

UNIVERSIDADE FEDERAL DO PARANÁ

RICARDO DALLA COSTA

**GENÉTICA DE POPULAÇÕES E MICROEVOLUÇÃO DO GENE *KIR3DL1/S1* NA
ESPÉCIE HUMANA**

CURITIBA

2011

RICARDO DALLA COSTA

**GENÉTICA DE POPULAÇÕES E MICROEVOLUÇÃO DO GENE *KIR3DL1/S1* NA
ESPÉCIE HUMANA**

Dissertação apresentada como requisito parcial para a obtenção do título de Mestre em Ciências Biológicas, na área de concentração de Genética, ao Curso de Pos-Graduação em Genética da Universidade Federal do Paraná

Orientadora: Profa. Dra. Maria Luiza Petzl-Erler.

CURITIBA

2011

Universidade Federal do Paraná
Sistema de Bibliotecas

Dalla-Costa, Ricardo

**Genética de Populações e Microevolução do Gene
KIR3DL1/S1 na Espécie Humana.**

Curitiba, 2011. 151p.

Orientadora: Profa. Dra. Maria Luiza Petzl-Erler

Dissertação (Mestrado) – Universidade Federal do Paraná (UFPR),
Setor de Ciências Biológicas. Programa de Pós-Graduação em
Genética.

1. *KIR*. 2. Polimorfismo genético. 3. Ameríndios. 4. Genética de
Populações.

Aqueles que acreditam na beleza
dos seus sonhos.

AGRADECIMENTOS

Agradeço,

De modo muito especial aos meus pais, aos meus irmãos e a minha avó Elvira, pelos anos de apoio e confiança. A maneira mais digna de eu agradecer por tudo que vocês me deram é procurar ser, a cada dia, um pouquinho melhor. Eu ainda espero vos encher de muito orgulho.

A Li, pela ajuda e apoio sem fim na conclusão deste trabalho, mas acima de tudo pela amizade e pelos sábios conselhos... que eu, teimoso como sempre, insisto em não ouvir.

A minha melhor amiga, Livia, que apesar de ter me trocado por Paris, vive em meu coração.

A Mah, pela linda amizade construída nesses últimos tempos e por ter me mostrado que o aMor existe.

Ao Vi, pelo carinho e atenção. Desculpa ser tão chato com você!

As amigas Giovana, Vanessa, Roseli e Ray pelos lindos dias vividos desde os tempos da faculdade. Sem vocês não teria sido fácil viver longe de casa. Obrigado por me suportarem!

Aos amigos do laboratório pelas risadas diárias, pelos problemas compartilhados, pelas discussões com as quais aprendo tanto.

Em especial à Sheyla, pelas muitas PCRs e reações de sequenciamento, que tornaram esse trabalho possível.

Ao Bruno, por enfrentar comigo esses dois anos de muito esforço e encarar o desafio de ser orientado pela professora Maria Luíza (risos).

A Áurea, pelos chocolates, pelos cafés, pelo carinho. Obrigado.

A professora Roseli Wassem pela ajuda e apoio irrestritos.

A todos, que de uma forma ou de outra, fizeram meus dias mais felizes nessa longa jornada de exaustivos dias de trabalho.

Por fim à professora Maria Luíza pela dedicação e competência com que orienta seus alunos. Obrigado pela oportunidade e confiança.

“Surpreender-se é sempre uma boa escolha!”

SUMÁRIO

| | |
|--|------------|
| 1 INTRODUÇÃO | 15 |
| 2 REVISÃO BIBLIOGRÁFICA..... | 17 |
| 2.1 CÉLULAS NK, FUNÇÕES E RECEPTORES DE MEMBRANA..... | 17 |
| 2.1.1 Genes que codificam os receptores das células NK | 20 |
| 2.1.1.1 Imunogenética dos receptores KIR..... | 25 |
| 2.1.1.2 Aspectos evolutivos dos genes <i>KIR</i> | 28 |
| 2.1.1.3 <i>KIR3DL1</i> e <i>KIR3DS1</i> | 35 |
| 2.2 POPULAÇÕES HUMANAS | 41 |
| 2.2.1 Populações Ameríndias | 41 |
| 2.2.1.1 As populações indígenas Kaingang e Guaraní | 43 |
| 2.2.2 População Brasileira | 46 |
| 3 OBJETIVOS E JUSTIFICATIVAS | 48 |
| 4 CAPÍTULO I – ARTIGO: High Polymorphism of KIR3DL1/S1 gene in Amerindians and Evidences of Natural Selection in Functional Domains | 49 |
| 5 RESULTADOS | 73 |
| 5.1 METODOLOGIA DE TIPAGEM E MÉTODOS DE INFERÊNCIA HAPLOTÍPICA | 73 |
| 5.2 DIVERSIDADE ALÉLICA E GENOTÍPICA | 75 |
| 5.3 DIVERSIDADE ESTRUTURAL E FUNCIONAL DO PRODUTO GÊNICO | 81 |
| 5.5 DIVERSIDADE NUCLEOTÍDICA E TESTES DE NEUTRALIDADE SELETIVA | 89 |
| 5.6 DISTÂNCIAS GENÉTICAS ENTRE AS POPULAÇÕES | 94 |
| 5.7 DESEQUILÍBRIO DE LIGAÇÃO | 99 |
| 6 DISCUSSÃO | 111 |
| 6.1 DIVERSIDADE GENÉTICA DO LOCO <i>KIR3DL1/S1</i> | 111 |
| 6.2 DIVERSIDADE GENÉTICA DO LOCO <i>KIR3DL1/S1</i> : EVOLUÇÃO NEUTRA? | 117 |
| 6.3 DIVERSIDADE FUNCIONAL DE RECEPTORES <i>KIR3DL1/S1</i> : SELEÇÃO BALANCEADORA? | 119 |
| 6.4 <i>KIR</i> E <i>HLA</i> : EVIDÊNCIAS DE COEVOLUÇÃO? | 120 |
| 6.5 DINÂMICA EVOLUTIVA DOS ALELOS <i>KIR3DL1/S1</i> | 120 |

| | |
|--|-----|
| 6.6 DESEQUILÍBRIO DE LIGAÇÃO | 123 |
| 7 CONCLUSÕES..... | 124 |
| 8 REFERÊNCIAS BIBLIOGRÁFICAS..... | 126 |
| 9 APÊNDICE: MATERIAIS E MÉTODOS | 141 |
| 9.1 CARACTERIZAÇÃO DAS AMOSTRAS | 141 |
| 9.2 AMPLIFICAÇÃO E SEQUENCIAMENTO | 144 |
| 9.3 ANÁLISE DOS DADOS | 149 |
| 9.3.1 Diversidade alélica, haplotípica e genotípica..... | 149 |
| 9.3.2 Análises evolutivas | 150 |

LISTA DE ABREVIATURAS

DNA – *Deoxiribonucleic Acid*

NK – *Natural Killer*

KIR – *Killer-cell Immunoglobulin-like Receptor*

MHC – *Major Histocompatibility Complex*

HLA – *Human Leukocyte Antigen*

IL – *Interleucina*

NKC – *Natural Killer Complex*

LRC – *Leukocyte Receptor Complex*

PCR – *Polymerase Chain Reaction*

dbLRC – *Leukocyte Receptor Complex database*

SNP – *Single Nucleotide Polymorphism*

AFRI – *Africanos*

EURO-BR – *Eurobrasileiros*

ASI-BR – *Brasileiros descendentes de Asiáticos*

GKW – *Guarani Kaiowá*

GND – *Guarani Nandeva*

GRC – *Guarani M'byá*

KIV – *Kaingang de Ivaí*

KRC – *Kaingang de Rio das Cobras*

ÍNDICE DE TABELAS

| | |
|---|-----|
| TABELA 1. Frequências dos genes <i>KIR3DL1</i> e <i>KIR3DS1</i> nas populações estudadas neste trabalho..... | 36 |
| TABELA 2. Comparação entre as frequências haplotípicas (alélicas) entre os diferentes métodos de inferência (2n=134 cromossomos)..... | 74 |
| TABELA 3. Frequências alélicas para o gene <i>KIR3DL1/S1</i> observadas em cada população deste estudo..... | 76 |
| TABELA 4. Número de alelos e heterozigosidade nas diferentes populações..... | 79 |
| TABELA 5. Análise da variância molecular (AMOVA) para o gene <i>KIR3DL1/S1</i> nas populações deste estudo..... | 80 |
| TABELA 6. Frequência dos motivos de interação com HLA nas diferentes populações..... | 83 |
| TABELA 7. Teste exato de diferenciação populacional baseado nas frequências dos motivos de ligação à HLA..... | 83 |
| TABELA 8. Número de substituições sinônimas e não sinônimas nos diferentes domínios do gene <i>KIR3DL1/S1</i> | 86 |
| TABELA 9. Proporção dN/dS (teste dN=dS) na sequência completa do gene <i>KIR 3DL1/S1</i> e nos exons 3, 4 e 5, que codificam os domínios extracelulares..... | 87 |
| TABELA 10. Frequências dos epítomos <i>Bw4</i> e <i>Bw6</i> e alelos <i>KR3DL1</i> e <i>3DS1</i> nas populações deste estudo..... | 89 |
| TABELA 11. Diversidade nucleotídica observada nas diferentes populações..... | 90 |
| TABELA 12. Parâmetros moleculares e testes de neutralidade seletiva para <i>3DL1/S1</i> , nas populações deste estudo..... | 92 |
| TABELA 13. Parâmetros moleculares e testes de neutralidade seletiva para os alelos <i>3DL1</i> nas populações deste estudo..... | 93 |
| TABELA 14. Teste D de Tajima para as regiões que compreendem os resíduos de interação com as moléculas HLA nos exons 4 e 5..... | 93 |
| TABELA 15. Distâncias genéticas entre as populações deste estudo..... | 94 |
| TABELA 16. Distâncias genética entre as populações deste estudo e outras populações.. | 96 |
| TABELA 17. Oligonucleotídeos iniciadores para as reações de amplificação e sequenciamento do gene <i>KIR3DL1/S1</i> | 145 |

ÍNDICE DE FIGURAS

| | |
|--|----|
| FIGURA 1. Regulação da atividade das células <i>NK</i> através da interação entre os receptores da sua superfície celular e os ligantes nas células alvo.. | 19 |
| FIGURA 2. Localização e organização genômica dos genes dos receptores das células <i>NK</i> humanas, pertencentes aos complexos <i>NKC</i> e <i>LRC</i> | 20 |
| FIGURA 4. Localização e organização genômica dos genes <i>KIR</i> | 22 |
| FIGURA 5. Estrutura dos receptores <i>KIR</i> | 23 |
| FIGURA 6. Modelo esquemático dos genes <i>KIR</i> | 24 |
| FIGURA 7. Origem dos genes <i>KIR</i> em primatas..... | 29 |
| FIGURA 8. Esquema hipotético para o surgimento dos genes <i>KIR</i> a partir do gene <i>KIR3DL</i> ancestral..... | 29 |
| FIGURA 9. Representação esquemática da região intrônica dos genes <i>KIR</i> , mostrando a presença de retroelementos e sequências <i>ALU</i> | 32 |
| FIGURA 10. Sequência de aminoácidos da proteína codificada pelo alelo <i>KIR3DL1*015</i> ... | 39 |
| FIGURA 11. Famílias, línguas e dialetos pertencentes ao tronco linguístico MACRO-JÊ. | 44 |
| FIGURA 12. Famílias, línguas e dialetos pertencentes ao tronco linguístico TUPI..... | 45 |
| FIGURA 13. Rede mutacional dos alelos <i>KIR3DL1/S1</i> | 78 |
| FIGURA 14. Cromatogramas mostrando as sequências identificadas para os novos alelos.. | 81 |
| FIGURA 15. Sequências de aminoácidos das proteínas codificadas pelos alelos <i>KIR3DL1/S1</i> identificados neste estudo. | 82 |
| FIGURA 16. Motivos de interação com as moléculas HLA formados por sítios nos domínios extracelulares D1 e D2 dos receptores <i>KIR3DL1/S1</i> | 82 |
| FIGURA 17. Rede mutacional dos motivos de ligação às moléculas HLA dos alelos <i>KIR3DL1/S1</i> identificados neste estudo..... | 84 |
| FIGURA 18. Regiões variáveis dos exons 3 (D0), 4 (D1) e 5 (D2) dos alelos conhecidos do gene <i>KIR3DL1/S1</i> . Em cinza são mostradas as novas variantes identificadas neste estudo. | 85 |
| FIGURA 19. Valores globais de <i>Fst</i> para os sítios polimórficos encontrados ao longo das sequências dos alelos identificados neste estudo..... | 88 |
| FIGURA 20. Relação entre as populações deste estudo, com base nas frequências alélicas do gene <i>KIR3DL1/S1</i> | 95 |

| | |
|---|-----|
| FIGURA 21. Relação entre as populações deste estudo (*) e outras populações da literatura, com base nas frequências alélicas do gene <i>KIR3DL1/S1</i> | 97 |
| FIGURA 22. Relação entre as populações deste estudo, com base nas frequências dos motivos de ligação às moléculas HLA dos alelos <i>KIR3DL1/S1</i> identificados nestas populações. | 98 |
| FIGURA 23. Padrão do desequilíbrio de ligação entre os 43 SNPs ao longo da sequência dos alelos identificados neste estudo na população africana (AFRI)..... | 100 |
| FIGURA 24. Padrão do desequilíbrio de ligação entre os 43 SNPs ao longo da sequência dos alelos identificados neste estudo na população eurobrasileira (EURO-BR)..... | 101 |
| FIGURA 25. Padrão do desequilíbrio de ligação entre os 43 SNPs ao longo da sequência dos alelos identificados neste estudo na população descendente de asiáticos (ASI-BR)... | 102 |
| FIGURA 26. Padrão do desequilíbrio de ligação entre os 43 SNPs ao longo da sequência dos alelos identificados neste estudo na população GKW. | 103 |
| FIGURA 27. Padrão do desequilíbrio de ligação entre os 43 SNPs ao longo da sequência dos alelos identificados neste estudo na população GND..... | 104 |
| FIGURA 28. Padrão do desequilíbrio de ligação entre os 43 SNPs ao longo da sequência dos alelos identificados neste estudo na população GRC..... | 105 |
| FIGURA 29. Padrão do desequilíbrio de ligação entre os 43 SNPs ao longo da sequência dos alelos identificados neste estudo na população KIV..... | 106 |
| FIGURA 30. Padrão do desequilíbrio de ligação entre os 43 SNPs ao longo da sequência dos alelos identificados neste estudo na população KRC. | 107 |
| FIGURA 31. Relação filogenética sugerida para os alelos <i>KIR3DL1/S1</i> e os novos alelos identificados neste estudo..... | 109 |
| FIGURA 32. Mapa da América do Sul, indicando a localização das populações analisadas neste estudo. | 142 |
| FIGURA 33. Mapa da África. | 143 |

RESUMO

As células *NK* são importantes para a defesa do organismo. O controle da sua atividade ocorre pelo balanço entre sinais ativadores e inibidores que são fornecidos por uma variedade de receptores presentes na superfície destas células. O gene *KIR3DL1/S1* possui grande diversidade alélica e codifica um desses receptores, que reconhece moléculas HLA com epítipo Bw4. O estudo de genes do sistema imune em diferentes populações é particularmente interessante para a genética de populações, pois muitos desses genes estão sujeitos a seleção natural acentuada, especialmente seleção positiva, flutuante e balanceadora, resultando em elevada taxa evolutiva, com importantes implicações na diversidade populacional. Neste trabalho, a diversidade do gene *KIR3DL1/S1* foi avaliada em amostras populacionais pertencentes às tribos indígenas Guarani e Kaingang dos estados do Mato Grosso do Sul e Paraná, além das populações africana, eurobrasileira e oriental, através da técnica de sequenciamento de DNA. A dinâmica evolutiva deste gene e sua coevolução com os ligantes HLA também foi avaliada nas populações deste estudo. Todas as populações estudadas exibem uma elevada diversidade para o gene *KIR3DL1/S1*, tendo sido identificados 25 alelos, incluindo quatro ainda não descritos, no conjunto total de indivíduos ($n=451$). Essa diversidade é próxima a observada para o sistema HLA, uma das regiões mais polimórficas do genoma humano. Além disso, as frequências alélicas diferem significativamente entre todas as populações, até mesmo entre as duas populações indígenas Kaingang, que são, historicamente e geneticamente bastante próximas. Essas diferenças resultam da ação de diferentes fatores evolutivos, especialmente do efeito fundador e da deriva genética, nas populações ameríndias. O fluxo gênico recente com não indígenas, já verificado em estudos anteriores, também pode ser verificado para o gene *KIR3DL1/S1* e contribui para aumento das diferenças genéticas entre os ameríndios. Apesar da grande diferenciação, as populações ameríndias são mais similares entre si e com os descendentes de asiáticos, o que está diretamente relacionado com a sua história evolutiva. O alelo *3DS1*01301* é o mais frequente em todas as populações indígenas e sua elevada frequência pode estar relacionada à importância dos genes *KIR* ativadores durante a evolução dessas populações. As maiores distâncias genéticas são observadas entre os indígenas e os africanos que, por sua vez, se aproximam mais dos eurodescendentes, provavelmente devido a contribuição africana na constituição genética da população eurobrasileira, como demonstrado por estudos anteriores. Os testes de neutralidade seletiva para a sequência completa, e para os exons separadamente, dos alelos identificados nas populações do presente estudo, não permitem rejeitar a hipótese de evolução neutra para este gene. Entretanto, quando analisados apenas os códons que codificam os motivos de interação com HLA, são observados indícios de seleção positiva e/ou balanceadora atuando sobre a diversificação desses domínios funcionais. A correlação positiva entre *3DL1* e a distribuição das moléculas HLA com epítipo Bw4 nessas populações também corrobora a hipótese de coevolução entre *KIR* e *HLA*.

Palavras chave: *KIR*; Polimorfismo Genético; Ameríndios, Genética de Populações.

ABSTRACT

The killer-cell immunoglobulin-like receptor (KIR) family is an important group of NK cell receptors that are expressed in a variegated manner on the NK cell population of a given individual. The interaction between KIR and HLA ligands can result in NK cell inhibition or activation and plays an important role in control of immune responses. The *KIR* genes are among the most polymorphic genes of the human genome and they have evolved rapidly, in a species-specific manner. At present, 84 alleles are assigned to the *KIR3DL1/S1* locus, and its products interact with the Bw4 motif of HLA-B and -A molecules. We analyzed 451 individuals, including Brazilians of predominantly European, of mixed but predominantly African and European, and of Asian ancestry, as well as five Amerindian populations from the Guarani and Kaingang groups, and an African population sample. Twenty five alleles were identified over all populations, including four previously unknown alleles. Despite the lowest genetic diversity in Amerindians, these populations presented high number of alleles and a considerable genetic diversity for the *KIR3DL1/S1* gene. Among the Amerindian populations analyzed, Guarani Nandeva had the highest gene flow from Europeans and Africans and accordingly had the highest genetic diversity. Instead, the Guarani M'byá had the lowest diversity. All populations differed greatly for allele frequencies. *KIR3DS1*01301* was the most common allele in Amerindians, accounting for up to 53% of the Guarani M'byá alleles. Conversely, the lineage *3DS1* was rare in Africans, as reported by others. Molecular analyzes showed an excess of non-synonymous substitutions in exons 4 and 5, at sites predicted to interact with HLA molecules but not in the other regions analyzed. In conclusion, these results showed that *KIR3DL1/S1* is highly polymorphic in all populations analyzed and there are evidences of coevolution between *KIR* and *HLA* loci.

Keywords: *KIR*; genetic polymorphism; Amerindians, population genetics

1 INTRODUÇÃO

O estudo da diversidade genética em diferentes populações auxilia o entendimento da evolução humana, da origem e da história demográfica das populações, além de permitir a identificação da ação de fatores evolutivos em sequências genômicas e da base genética de doenças.

As populações podem diferir quanto às frequências alélicas. Alelos muito comuns em um grupo podem ser raros ou mesmo ausentes em outro grupo, especialmente quando as populações analisadas são geneticamente isoladas. A identificação de genes cujas frequências alélicas diferem entre populações permite que os alelos destes genes sejam usados como marcadores de ancestralidade. Desse modo, o estudo de genes e de sua variabilidade, pode ser extremamente útil na reconstituição da história evolutiva dos grupos humanos.

As populações indígenas são particularmente interessantes para os estudos populacionais, pois são populações isoladas, nas quais o fluxo gênico recente é bem menos intenso quando comparado ao que vem ocorrendo historicamente entre as populações urbanas. Estudos genéticos de populações ameríndias, aliados a abordagens arqueológicas e linguísticas, levaram à elaboração de hipóteses sobre o povoamento das Américas. Trabalhos que investiguem a diversidade genética de populações indígenas, bem como de outras populações, podem contribuir significativamente para a elucidação da história das populações humanas.

A imunidade inata constitui a primeira linha de defesa do organismo contra infecções, o desenvolvimento de tumores e agentes estranhos ao organismo. As células *NK* (*Natural Killer*) são uma importante subpopulação de linfócitos atuantes nesta fase da resposta imunológica. Além disso, estas células também são importantes no desencadeamento das respostas imunológicas adaptativas. O controle da ação das células *NK* se dá através de receptores de membrana, entre os quais estão os receptores KIR (*Killer-Cell Immunoglobulin like-receptors*), que reconhecem moléculas MHC (*Major Histocompatibility Complex*) de classe I, expressas pela maioria das células normais do organismo. Os receptores KIR são codificados por uma família de genes polimórficos, entre os quais está o gene *KIR3DL1/S1*, que possui grande importância na regulação da atividade das células *NK*.

A caracterização de populações para genes que codificam moléculas importantes durante as respostas imunológicas tem a aplicação adicional de servir como referência para

estudos de associação a doenças, pois estes genes podem ter um papel importante na patogênese e seus polimorfismos podem estar associados à maior ou menor susceptibilidade ao desenvolvimento destas doenças.

Neste trabalho, a diversidade alélica do gene *KIR3DL1/S1* foi investigada nas populações ameríndias Guarani e Kaingang, na população curitibana, de origem predominantemente européia e africana, em uma amostra populacional de ancestralidade oriental e em uma população africana, com o intuito de entendermos um pouco mais acerca das relações evolutivas entre as diferentes populações e dos fatores evolutivos que atuam sobre as sequências genômicas.

2 REVISÃO BIBLIOGRÁFICA

2.1 CÉLULAS *NK*, FUNÇÕES E RECEPTORES DE MEMBRANA

A imunidade inata constitui a primeira linha de defesa do organismo contra infecções, o desenvolvimento de tumores e agentes estranhos ao organismo. A resposta inflamatória resultante da invasão do organismo por algum patógeno recruta moléculas e células efectoras do sistema imune que combatem sua propagação. Somente quando essa barreira é vencida, a doença infecciosa instala-se e uma resposta imunológica adaptativa é desencadeada.

As células *NK* (*natural killers* ou assassinas naturais) constituem uma importante população de linfócitos que atua no combate às células infectadas ou malignamente transformadas, representando cerca de 10-15% dos linfócitos circulantes e uma expressiva porcentagem dos linfócitos teciduais (COOPER, FEHNIGER e CALIGIURI, 2001; BAXTER e SMYTH, 2002;).

Através da liberação de grânulos contendo granzimas e perforinas, as células *NK* possuem atividade citotóxica direta, induzindo a apoptose das células alvo. As células *NK* também secretam moléculas imunorreguladoras, como citocinas e quimiocinas, que recrutam outras células de defesa do organismo para os sítios inflamatórios. Apesar de não exercerem ação antígeno específica, as células *NK* também são importantes nas fases iniciais das respostas imunológicas adaptativas, (COOPER, FEHNIGER e CALIGIURI, 2001; ROBERTSON, 2002; VIVIER, NUNÈS e VÉLY, 2004; CALIGIURI, 2008). A interação de receptores das células *NK* com moléculas da superfície das células dendríticas auxilia na maturação destas células, que por sua vez estão diretamente relacionadas à ativação dos linfócitos T e à diferenciação e proliferação dos linfócitos B. A produção de IFN- γ (Interferon-gamma) e de outras citocinas pelas células *NK* também atua na ativação dos macrófagos e de outras células apresentadoras de antígenos (APCs) e na ativação e diferenciação dos linfócitos T em células efectoras (COOPER, FEHNIGER e CALIGIURI, 2001; ROBERTSON, 2002; di SANTO, 2006; CALIGIURI, 2007).

Além de atuarem na primeira linha de defesa do organismo contra células infectadas e malignamente transformadas e no desencadeamento das respostas imunológicas adaptativas, as células *NK* também desempenham um importante papel na reprodução. Durante o estágio de implantação do embrião no útero materno, as células *NK* uterinas são

ativadas, através da interação com moléculas da superfície do trofoblasto, secretando citocinas que atuam na remodelação do epitélio arterial materno, permitindo a invasão do trofoblasto e formação adequada da placenta (PARHAM, 2004; TROWSDALE e MOFFETT, 2008).

Embora a função efetora das células *NK* seja extremamente importante no combate às infecções e às células tumorais, a regulação da sua atividade é essencial para a manutenção da tolerância aos antígenos próprios e para o sucesso gestacional. Muitos estudos relacionam a atividade das células *NK* com o desenvolvimento de doenças auto-imunes. Evidências sugerem a existência de associação entre a diminuição no número e na atividade das células *NK* e as doenças auto-imunes esclerose múltipla, diabetes tipo I e lúpus eritematoso sistêmico (BAXTER e SMYTH, 2002; PARHAM 2005b; LORENTZEN et al., 2009). Além disso, um controle insuficiente da ativação das células *NK* decíduais pode favorecer o abortamento espontâneo (VARLA-LEFTHERIOTI et al., 2003)

As células *NK* são derivadas de progenitores hematopoiéticos na medula óssea e sua diferenciação e desenvolvimento podem ocorrer tanto na medula, quanto nos órgãos linfóides periféricos (COOPER, FEHNIGER e CALIGIURI, 2001; CALIGIURI, 2008). A regulação das células *NK* ocorre principalmente nas fases iniciais do seu desenvolvimento, diferenciação e ativação, e envolve múltiplos mecanismos. As primeiras evidências sugerem que a expressão de pelo menos um receptor inibidor, capaz de reconhecer moléculas MHC (*Major Histocompatibility Complex*) de classe I próprias, durante o seu desenvolvimento, seria fundamental para a tolerância aos constituintes próprios e para a ação das células *NK* contra infecções e transformações malignas (RAULET, VANCE e McMAHON, 2001). As células que expressam esse receptor são denominadas competentes ou “licenciadas” (PARHAM, 2006; YOKOYAMA e KIM, 2006; ANFOSSI et al., 2006; KULKARNI, MARTIN e CARRINGTON, 2008; YAWATA et al., 2008; KIM et al., 2008). Por outro lado, estudos recentes demonstram que as células *NK* que não expressam nenhum receptor inibidor MHC-específico continuam tolerantes aos antígenos próprios, porém tornam-se anérgicas e irresponsivas (COOLEY et al., 2007).

A expressão progressiva de receptores *KIR* (*Killer Cell Immunoglobulin like-receptors*) inibidores que reconhecem moléculas MHC de classe I próprias está relacionada com um aumento da capacidade efetora das células *NK* (YU et al., 2007). Além disso, o processo de licenciamento pode contribuir para a heterogeneidade nos níveis de ativação das células *NK*, observada entre diferentes indivíduos e até mesmo entre diferentes subpopulações de células *NK* de um mesmo indivíduo, uma vez que os genes *KIR* são

expressos de modo aleatório e determinadas interações KIR:HLA (*Human Leukocyte Antigen*) têm maior efeito sobre as células *NK* (KIM et al., 2008).

Embora a seleção ontogenética seja o principal mecanismo de regulação das células *NK*, nos tecidos periféricos os receptores ativadores e inibidores são bastante importantes para a regulação destas células (Figura 1). Nestes tecidos o delicado equilíbrio da atividade das células *NK* ocorre através da integração de inúmeros sinais produzidos pela ligação de moléculas sinalizadoras (por exemplo, IL-2, -10, -12, -15 e o Fator de Necrose Tumoral TNF- α/β) aos receptores ativadores e inibidores expressos na superfície das células *NK* (BAXTER e SMYTH, 2002; ROBERTSON, 2002 DRAGHI et al., 2005; di SANTO, 2006; ANFOSSI et al., 2006; CALIGIURI, 2007; PARHAM et al., 2010). A ativação ou inibição resulta do balanço entre os sinais ativadores e inibidores gerados a partir destes receptores (JANEWAY et al., 2002; VIVIER, NUNÈS e VÉLY, 2004).

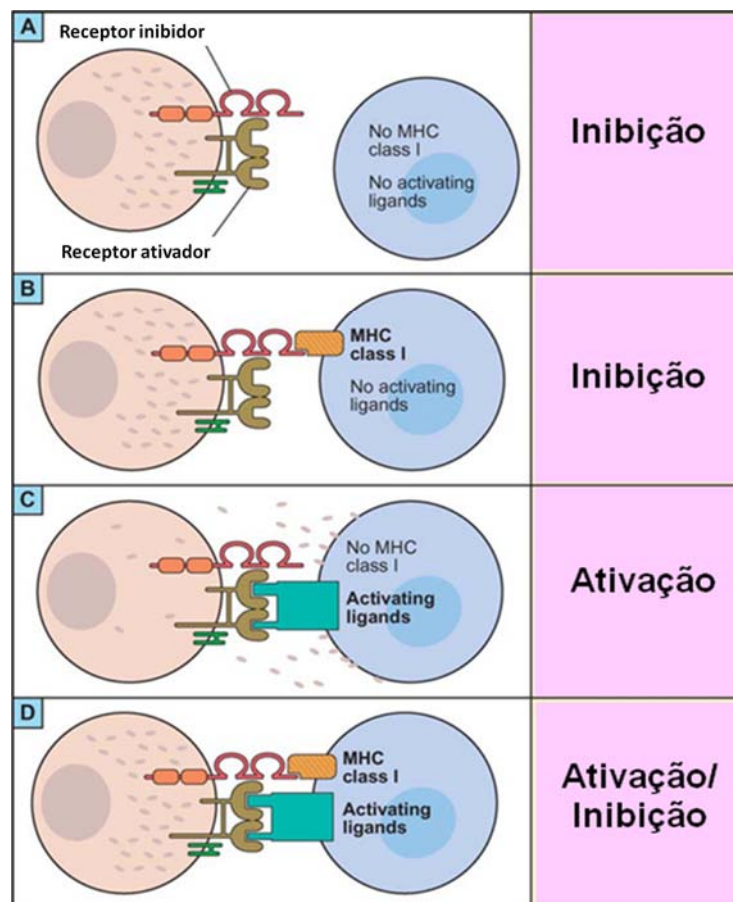


FIGURA 1. Regulação da atividade das células *NK* através da interação entre os receptores da sua superfície celular e os ligantes nas células alvo. **Fonte:** di SANTO, 2006.

2.1.1 Genes que codificam os receptores das células NK

Em humanos, os genes que codificam os receptores das células *NK* estão agrupados principalmente em dois cromossomos distintos (Figura 2). No cromossomo 12 está localizado o complexo *NKC* (*Natural Killer Complex*), no qual são encontrados os genes que codificam receptores semelhantes à lectina C, como *CD94* e *NKG2*. O outro complexo, conhecido como *Leukocyte Receptor Complex (LRC)*, está localizado no cromossomo 19 e seus genes codificam moléculas relacionadas à superfamília das imunoglobulinas, como os receptores *KIR* (BROWN et al., 2000; MORETTA et al., 2000; PARHAM et al., 2010).

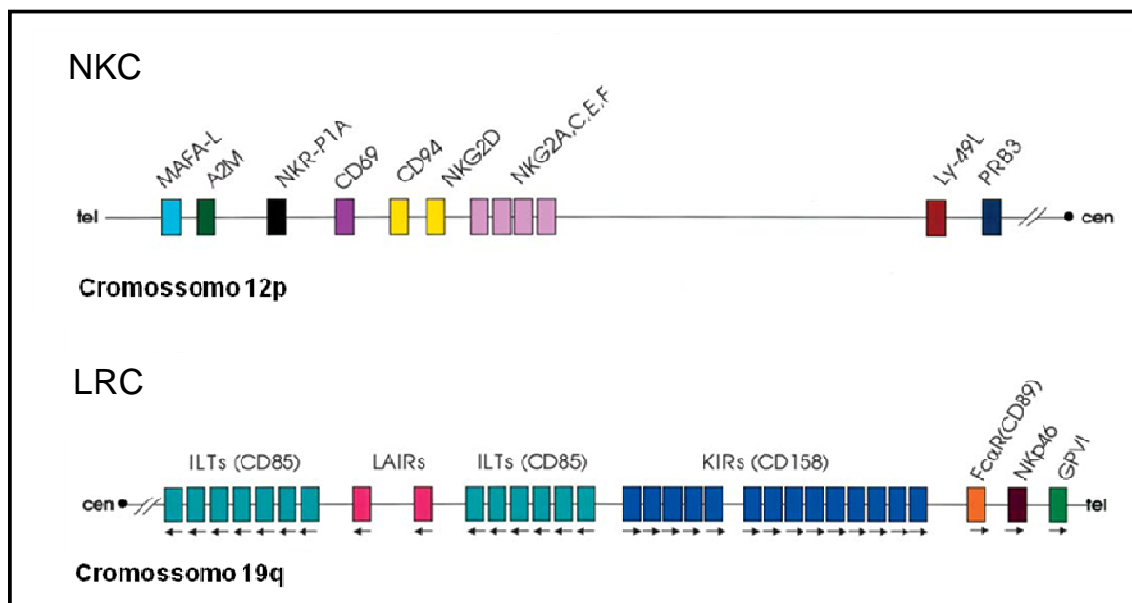


FIGURA 2. Localização e organização genômica dos genes dos receptores das células *NK* humanas, pertencentes aos complexos *NKC* e *LRC*. **Fonte:** HSU et al., 2002.

A composição, estrutura e organização dos complexos gênicos *NKC* e *LRC* é muito variável, mesmo entre espécies próximas, evidenciando que a evolução destes genes ocorreu de forma bastante acelerada e de modo espécie específico (PARHAM et al., 2010; MOESTA et al., 2009; KELLEY, WALTER e TROWSDALE, 2005; KAKOO et al., 2000). Além disso, alguns estudos demonstram a ocorrência de expansões e deleções gênicas nos complexos que codificam os receptores das células *NK* e nos genes que codificam seus respectivos ligantes, evidenciando processos coevolutivos entre estes genes em diferentes espécies (KHAKOO et al., 2000; GUETHLEIN et al., 2002; GUETHLEIN et al., 2007; HOLLENBACH et al., 2010; PARHAM et al., 2010).

As células *NK* expressam um conjunto diverso de receptores ativadores e inibidores em suas membranas, como por exemplo, os receptores KIR, CD94-NKG2 (*C-lectin-like killer inhibitory receptor*), ILTs (*Immunoglobulin-Like Transcript Receptors*) e *Natural Cytotoxicity Receptor* NKp30 e NKp44. Estes receptores reconhecem moléculas MHC de classe I próprias (HLA-A, -B, -C, -E e -G) e moléculas estruturalmente relacionadas ao MHC, como MIC-A (*MHC Class I Related Chain A*) e MIC-B (*MHC Class I Related Chain B*), além de outras moléculas (Figura 3A e 3B) (TROWSDALE, 2001; CALIGIURI, 2007; PARHAM et al., 2010). Embora sejam bem caracterizados os receptores inibidores das células *NK* e seus ligantes, para a maioria dos receptores ativadores, ainda não se conhecem quais são os ligantes. Cada célula *NK* expressa, aleatoriamente, um conjunto diferente de receptores ativadores e inibidores. Dessa forma, nem todas as células *NK* do organismo apresentam os mesmos receptores em sua membrana, podendo cada clone ser mais ou menos reativo a determinados patógenos ou células neoplásicas (FREUD e CALIGIURI, 2006).

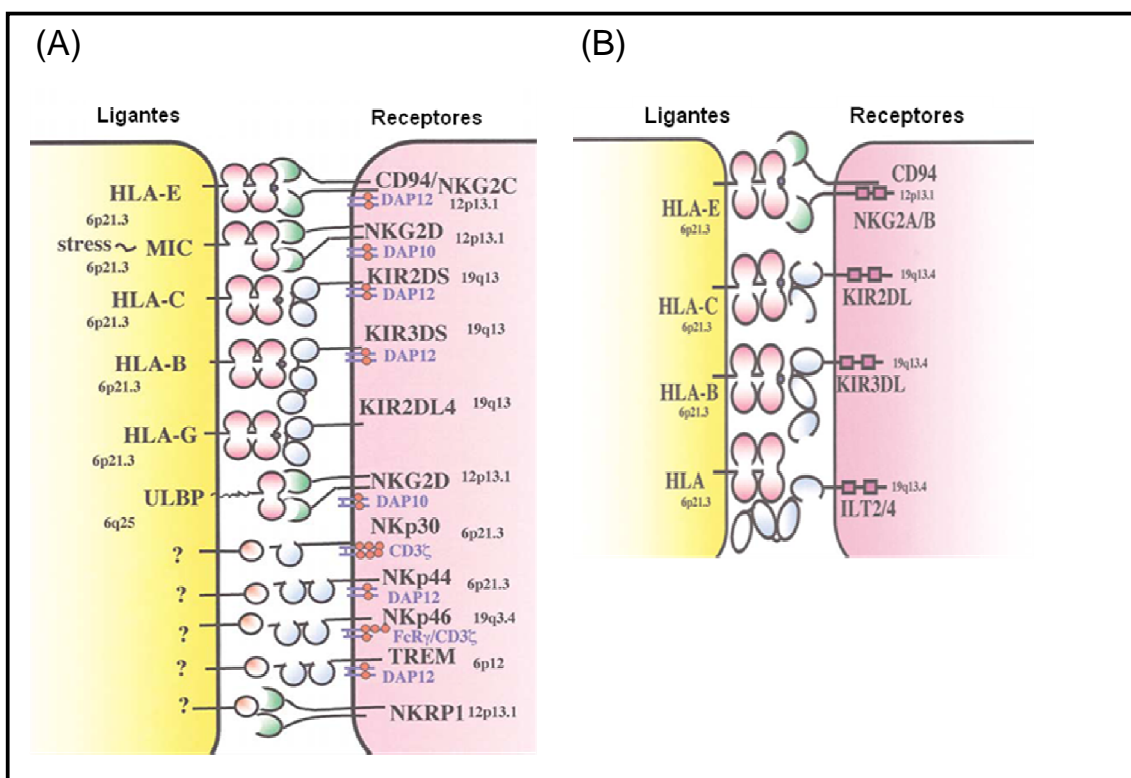


FIGURA 3. Receptores ativadores (A) e receptores inibidores (B) das células *NK* e seus ligantes conhecidos. **Fonte:** TROWSDALE, 2001.

Dentre os receptores que participam da regulação da atividade das células *NK*, estão os receptores KIR, que são expressos na superfície destas células e também em algumas subpopulações de linfócitos T (UHRBERG et al., 2001). Os receptores KIR pertencem à superfamília das imunoglobulinas e são codificados por uma família de genes polimórficos localizada no cromossomo 19 (19q13.4), no complexo *LRC* (Figura 4). São conhecidos 13 genes *KIR* e dois pseudogenes, dispostos em *tandem* em um segmento de aproximadamente 150 kb de DNA.

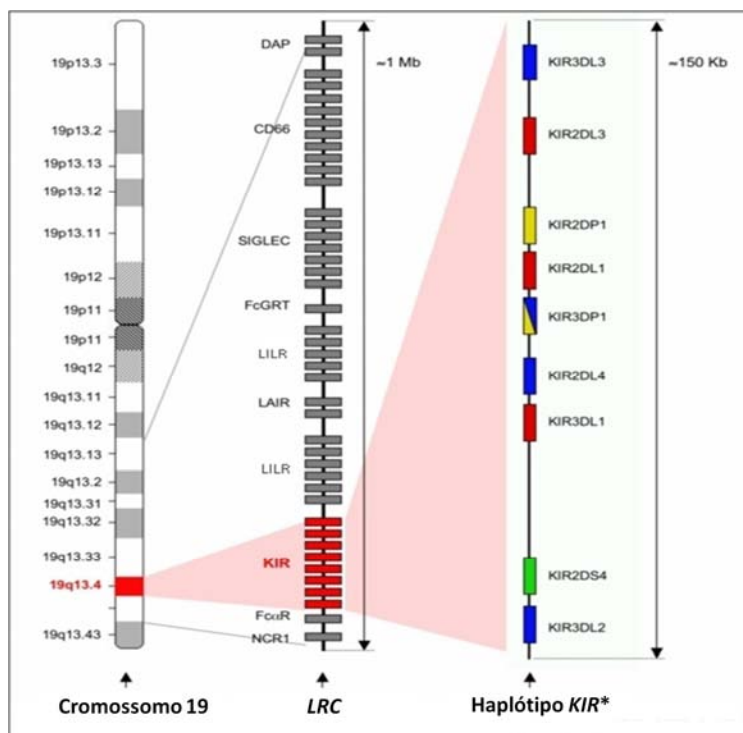


FIGURA 4. Localização e organização genômica dos genes *KIR*. Os genes *KIR* distribuem-se por cerca de 150 kb de DNA e estão localizados no braço curto do cromossomo 19 (19q13.4). *LRC* - *Leukocyte Receptor Complex*. * - Exemplo de um haplótipo dos genes *KIR* (nem todos os genes estão presentes nesse haplótipo). **Fonte:** IPD-KIR Database. Disponível em: <<http://www.ebi.ac.uk/ipd/kir/introduction.html>>. Acesso em: 30/07/2008.

A nomenclatura dos genes *KIR* é baseada na estrutura molecular de seus produtos. São denominados 2D os genes que codificam proteínas com dois domínios extracelulares e 3D os genes que codificam proteínas com três domínios. As letras L e S (do inglês *Long* e *Short*) referem-se ao tamanho da cauda citoplasmática, longa ou curta, respectivamente (Figura 5) (PARHAM, 2005a). Os receptores de cauda citoplasmática longa possuem motivos *ITIM* (imunoreceptor inibidor baseado em tirosina) e atuam como inibidores da transdução intracelular de sinais, interrompendo a cascata de ativação celular. Por outro

lado, os receptores de cauda citoplasmática curta possuem motivos *ITAM* (imunoreceptor ativador baseado em tirosina) e estão associados à molécula sinalizadora DAP12, atuando como ativadores das células *NK* (SALTER et al., 1997; WILSON et al., 2000). O gene *KIR2DL4* é uma exceção, pois atua tanto na transdução de sinais ativadores como na transdução de sinais inibidores (RAJAGOPALAN e LONG, 2001; YUSA, CATINA e CAMPBELL, 2001; BIASSONI et al., 2003).

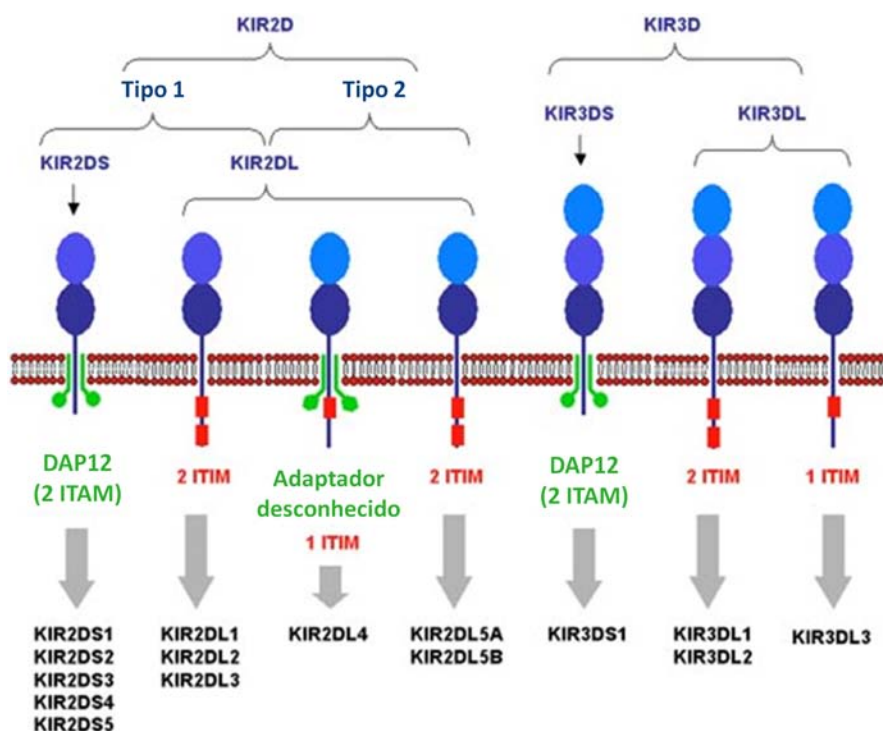


FIGURA 5. Estrutura dos receptores KIR. São mostrados os domínios extracelulares, os motivos *ITIM* e as moléculas adaptadoras DAP12. **Fonte:** IPD-KIR Database. Disponível em: <<http://www.ebi.ac.uk/ipd/kir/introduction.html>>. Acesso em: 30/07/2008.

Os genes *KIR* estendem-se por segmentos que variam entre 10 e 16 kb de DNA e, individualmente, estão organizados em nove exons que codificam os diferentes domínios funcionais da proteína (Figura 6). Os exons um e dois codificam o peptídeo sinal, enquanto que os exons três, quatro e cinco codificam os três domínios extracelulares semelhantes à imunoglobulina, D0, D1 e D2, respectivamente, que são responsáveis pelo reconhecimento dos ligantes. O exon seis codifica o segmento que une D2 a região transmembrana, que por sua vez é codificada pelo exon sete. Os exons oito e nove codificam o domínio citoplasmático da molécula, responsável pela transdução intracelular de sinais. Os receptores com dois domínios extracelulares são codificados por genes que possuem uma

região conhecida como “pseudoexon três”, que é excisada durante a maturação do RNA mensageiro (VILCHES e PARHAM, 2002).

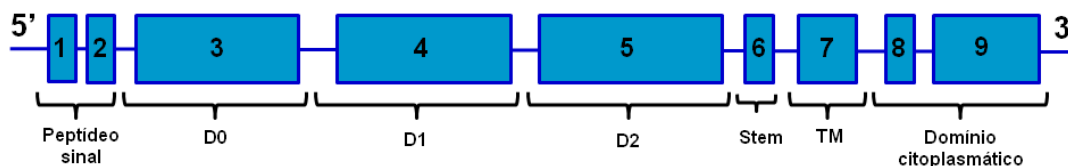


FIGURA 6. Modelo esquemático dos genes *KIR*. São mostrados os exons (1-9) e os domínios da proteína que são codificados por cada exon. D0-2 - Domínios extracelulares semelhantes à imunoglobulina; TM - Domínio transmembrana. **Fonte:** o AUTOR, 2009.

Os receptores KIR são classificados dentro de cinco linhagens, de acordo, principalmente, com a organização dos seus domínios extracelulares. Três destas linhagens são conservadas entre humanos e chimpanzés. A linhagem I é formada pelos genes *KIR2DL4* e *KIR2DL5*, que codificam os receptores que possuem os domínios extracelulares D0 e D2. A linhagem II agrupa os genes que codificam moléculas com três domínios extracelulares (D0, D1 e D2). Os receptores que possuem os domínios extracelulares D1 e D2, como por exemplo, *KIR2DS4*, constituem a terceira linhagem (VILCHES e PARHAM, 2002). As linhagens IV e V não são conservadas entre as espécies.

Os receptores KIR representam um importante repertório funcional para as células *NK* e, embora sejam conhecidos tanto receptores KIR ativadores como inibidores, a função inibidora parece ser preponderante para regulação das células *NK*. Estes receptores ligam principalmente moléculas *MHC* de classe I, porém para vários dos receptores KIR ainda não se conhecem os ligantes. Embora os genes *KIR* e os genes que codificam as moléculas *MHC* de classe I (*HLA* de classe I em humanos) segreguem independentemente, a estreita relação funcional entre estes genes sugere que a expressão de ambos esteja inter-relacionada. A interação KIR:HLA é fundamental na educação (ou licenciamento) das células *NK* (ANFOSSI et al., 2006; YAWATA et al., 2006; YOKOYAMA e SKIM, 2006; PARHAM et al., 2010).

As células *NK* podem expressar um conjunto variado de receptores KIR em sua superfície. Entre diferentes indivíduos, essa diversidade de receptores KIR é dada, principalmente, pelas diferentes combinações dos genes *KIR* nos haplótipos. Além disso, polimorfismos presentes principalmente na região promotora destes genes, onde se ligam fatores de transcrição, podem explicar por que, em um mesmo indivíduo, subpopulações de

células *NK* expressam diferentes conjuntos de receptores KIR (KELLEY, STEWART, BERGEN e TROWSDALE, 2003; WALTER e TROWSDALE, 2005; LI et al., 2008; THOMAS et al., 2008). Mecanismos epigenéticos, como a condensação da cromatina e a metilação, também podem atuar sobre o controle da expressão destes genes, inibindo sua expressão ou favorecendo a expressão de um alelo em detrimento de outros (CHAN et al., 2003; GAO et al., 2009; GATTAZZO et al., 2010).

Embora algumas evidências suportem a hipótese de que a expressão dos genes *KIR* seja independente da expressão das moléculas MHC de classe I (ANDERSSON et al., 2009), outros trabalhos mostram que a expressão destes genes pode ser influenciada pelo genótipo *HLA* (SHILLING et al., 2002; YAWATA et al., 2006; YU et al., 2007). Indivíduos que possuem o gene *KIR2DL1* e alotipos *HLA-C* do grupo Cw2 (com lisina na posição 80) apresentam células *NK* que expressam em maior nível o gene *KIR2DL1* quando comparadas com as que não possuem ligante Cw2. De modo semelhante, indivíduos que possuem ligantes *HLA* com epítipo Bw4 apresentam maiores níveis de expressão do gene *KIR3DL1* em comparação com indivíduos que não possuem moléculas *HLA-Bw4* (PARHAM, 2006; PARHAM et al., 2010). Além disso, os receptores KIR parecem agir sinergisticamente, induzindo a expressão de outros receptores KIR (ANDERSSON et al., 2009).

2.1.1.1 Imunogenética dos receptores KIR

Os genes *KIR* atuam epistaticamente no controle da atividade das células *NK*, e cada célula pode expressar um conjunto diferente de receptores KIR que possuem funções correspondentes (SINGLE et al., 2007). A expressão dos genes *KIR* é um dos principais mecanismos envolvidos na regulação da atividade das células *NK*, possibilitando a tolerância aos constituintes próprios do organismo e a ativação dos mecanismos de defesa contra a invasão do organismo por patógenos ou o surgimento de células neoplásicas (HSU et al., 2002).

Na espécie humana, os genes *KIR* formam haplótipos que são agrupados em dois haplogrupos, o haplogrupo A e o haplogrupo B. O haplogrupo A é representado por um único haplótipo, que contém sete genes, dos quais apenas *KIR2DS4* codifica um receptor ativador. Este gene possui um alelo não funcional que ocorre em frequências elevadas em todas as populações estudadas. Dessa forma, o haplótipo A tem um papel predominantemente inibidor da ação das células *NK*. O haplogrupo B é formado por um grande número de haplótipos, contendo um número variável de genes, entre os quais estão

presentes outros genes *KIR* ativadores, além de *KIR2DS4* (PARHAM et al., 2005b). Com exceção de *KIR2DL4*, que é expresso por todas as *NK*, os demais genes *KIR* são expressos em diferentes combinações pelas *NK*. Essa expressão diferencial entre as células *NK* de um mesmo indivíduo ocorre pela ação de mecanismos epigenéticos que regulam a expressão dos diferentes genes *KIR* (GARDNER, 2007).

Os receptores *KIR* inibidores reconhecem moléculas MHC de classe I, expressas pela grande maioria das células normais do organismo. Embora alguns receptores *KIR* ativadores também se liguem à moléculas do MHC, a maior parte dos receptores *KIR* ativadores não possui ligantes conhecidos. Na presença de sinais ativadores e inibidores a influência inibidora é predominante e a célula *NK* se encontra inativa, impedindo assim a morte de células normais. A transformação maligna das células ou a infecção por microrganismos patogênicos pode levar a expressão alterada das moléculas MHC de classe I, de modo que o sinal inibidor é perdido e as células *NK* são ativadas (LANIER, 2005). Esse modelo de ativação é conhecido como “*missing-self*” (falta do próprio), pois as células assassinas naturais agem quando não encontram as moléculas que conferem a identidade da célula própria, ao contrário dos linfócitos T que geralmente agem pelo reconhecimento de moléculas alteradas (“*altered-self*”) (LJUNGGREN e KARRE, 1990; KÄRRE, 2002).

Os receptores *KIR* ativadores possuem em seu domínio citoplasmático de um a três motivos *ITAM* (imunoreceptor ativador baseado em tirosina) e geralmente estão associados, não-covalentemente, à molécula DAP12, levando a produção citocinas e quimiocinas, como interferons e TNF (fator de necrose tumoral), que leva a completa ativação e diferenciação das células *NK*. Ao contrário, os receptores inibidores estão associados a motivos *ITIM* (imunoreceptor inibidor baseado em tirosina), que recrutam as proteínas SHP1 e SHP2 (proteínas tirosina-fosfatases) e atuam bloqueando a transdução dos sinais ativadores nas células *NK* (LJUNGGREN e KARRE, 1990; SALTER et al., 1997; WILSON et al., 2000; BAXTER e SMYTH, 2002; YUSA, CATINA e CAMPBELL, 2002; VIVIER, NUNÈS e VÉLY, 2004).

Embora seja considerável a importância da ativação das células *NK* para a imunidade inata, os receptores inibidores são fundamentais para o controle da atividade das células *NK* e para a manutenção da tolerância periférica. O elevado grau de polimorfismo tanto nos genes *KIR* quanto nos ligantes *HLA* e a segregação independente destes genes permite considerar que certas combinações *HLA* e *KIR* representam um fator de risco relevante na patogênese de doenças. Vários estudos recentes correlacionaram à incidência e a progressão de diversas doenças com determinados genes *KIR* em particular, entre as quais doenças infecciosas, doenças auto-imunes, doenças relacionadas à reprodução e o

câncer (PARHAM, 2005b; KULKARNI, MARTIN e CARRINGTON, 2008). De modo geral, genótipos *KIR:HLA* que possuem efeito ativador forte são benéficos no combate a infecções virais, mas podem representar um fator de risco relevante ao desenvolvimento de doenças auto-imunes (KULKARNI, MARTIN e CARRINGTON, 2008). A variação do conteúdo gênico e a diversidade alélica dos genes *KIR* e *HLA* podem afetar a função e o nível de especificidade e sensibilidade das células *NK* (KIM et al., 2007).

As diferentes combinações *KIR:HLA* também desempenham um importante papel na resposta de indivíduos transplantados com células tronco hematopoiéticas retiradas da medula óssea. O transplante entre indivíduos que não são completamente idênticos para os ligantes *HLA* parece facilitar a completa eliminação das células neoplásicas pelas células *NK* provenientes do doador em alguns tipos de leucemia (PARHAM, 2004; KELLEY e TROWSDALE, 2005; VELARDI, 2008). Ao contrário, a presença dos genes *KIR2DS1*, em conjunto com a ausência dos ligantes dos genes *KIR2DL2* e *2DL3* (*HLA-Cw1*), está associada a uma pior recuperação de pacientes com leucemia mielóide após transplante de medula (COOK et al., 2004).

Muitos pesquisadores investigaram a relação entre determinadas doenças e os receptores das células *NK*, particularmente dos receptores *KIR* e seus ligantes conhecidos. Estudos mostram que os genes *KIR2DS1* e *2DS2* estão associados a uma maior susceptibilidade à artrite psoriática caso haja ausência de ligantes *HLA-Cw* para os receptores inibidores *KIR2DL1* e/ou *2DL2/2DL3* (MARTIN et al., 2002). O diabetes tipo 1 foi associado ao aumento no número de genes *KIR* ativadores, particularmente com a presença *KIR2DS2* e *HLA-Cw* com asparagina na posição 80 (grupo *HLA-Cw1*), e a ausência de pares inibidores *KIR:HLA* (van der SLIK et al., 2003). De forma semelhante, a combinação entre *KIR2DS2* e *HLA-Cw1* está associada à vasculite reumatóide, que ocorre como complicação da artrite reumatóide (YEN et al., 2001). Recentemente, foi demonstrado que indivíduos portadores do gene *KIR2DL1* e *HLA-Cw2* possuem um menor risco para o desenvolvimento de esclerose múltipla (LORENTZEN et al., 2009) e diabetes tipo 1 (JOBIM et al., 2010). Outros estudos revelam a importância da interação de determinados receptores *KIR* e seus ligantes no combate a doenças infecciosas. Indivíduos homozigotos para a presença de *KIR2DL3* e ligantes do grupo *HLA-Cw1* possuem uma expressiva vantagem no combate ao vírus da hepatite C, em relação aos demais (KHAKOO et al. 2004). Determinados genótipos *KIR3DL1/S1* e alelos do gene *KIR2DS4* foram positivamente correlacionados com o desenvolvimento de malária (TANIGUCHI e KAWABATA, 2009). Além disso, nas infecções com o vírus HIV (*Human Immunodeficiency Virus*), a progressão da infecção é menor nos pacientes que possuem o receptor *KIR3DS1* e um ligante *HLA* com

epítopo Bw4 que contenha isoleucina na posição 80, demonstrando a interação epistática entre estas moléculas (MARTIN et al., 2002; MARTIN et al., 2007). Gaudieri e colaboradores (2005) demonstraram que indivíduos portadores do gene *KIR2DS2* apresentam uma acelerada progressão da infecção pelo vírus HIV em relação aos demais. A expressão anormal de alguns genes *KIR* e a ausência de determinados receptores KIR ativadores já foram associadas com o desenvolvimento de certas neoplasias (KULKARNI, MARTIN e CARRINGTON, 2008; BONAGURA et al., 2010; GATTAZZO et al., 2010). Estudos que relacionam genes *KIR* com doenças reprodutivas revelam que mulheres homozigotas para o haplogrupo A, cujos fetos sejam homozigotos para alótipos do grupo *HLA-C2*, apresentam maior risco no desenvolvimento da pré-eclampsia (HIBY et al., 2004).

No entanto, é importante salientar que essas associações de doenças a determinados genes *KIR* podem ser decorrentes de desequilíbrio de ligação com outros locos. Dessa forma, ainda permanecem desconhecidos os reais motivos que levam células *NK*, responsáveis pela defesa do organismo, a estarem associadas às doenças auto-imunes.

2.1.1.2 Aspectos evolutivos dos genes *KIR*

Análises filogenéticas das sequências dos genes *KIR* em diferentes espécies mostram que o primeiro passo na diversificação destes genes ocorreu há cerca de 136 milhões de anos, com a duplicação de um gene *KIR3D* ancestral, originando os genes *KIR3DL* e *KIR3DX* (Figura 7) (HUGHES, 2002; SAMBROOK et al., 2006; PARHAM et al., 2010). Nos primatas, a cerca de 40-58 milhões de anos, o gene *KIR3DL* sofreu ampla diversificação, dando origem aos diferentes genes *KIR* que conhecemos atualmente (Figura 8), enquanto que na maioria das demais espécies de mamíferos ele tornou-se não-funcional ou foi perdido (PARHAM et al., 2010). Mesmo entre espécies de primatas proximamente relacionadas, como chimpanzés e humanos, os genes *KIR* são bastante diversos, o que permite sugerir que estes genes evoluíram rapidamente e de modo espécie-específico (KHAKOO et al., 2000; HERSHBERGER et al., 2001; MORETTA e MORETTA, 2004; KELLEY, WALTER e TROWSDALE, 2005; SAMBROOK et al., 2005; HERSHBERGER et al., 2005; MOESTA et al., 2009;).

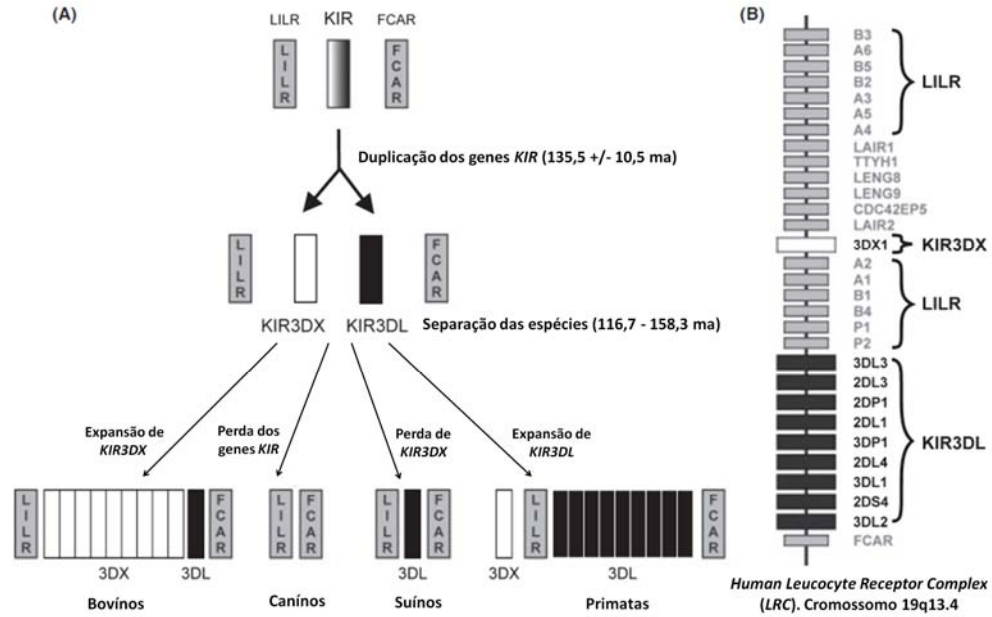


FIGURA 7. Origem dos genes *KIR* em primatas. (A) Duplicação do gene *KIR3D* ancestral, originando os genes *KIR3DX* e *KIR3DL*. Durante a evolução dos primatas, o gene *KIR3DX* permaneceu como uma única cópia, enquanto *KIR3DL* diversificou-se e originou os genes *KIR* conhecidos atualmente. (B) Organização dos genes humanos do complexo *LRC*, mostrando a localização do gene *KIR3DX*. Apenas alguns dos genes *KIR* são representados. **Fonte:** PARHAM et al., 2010.

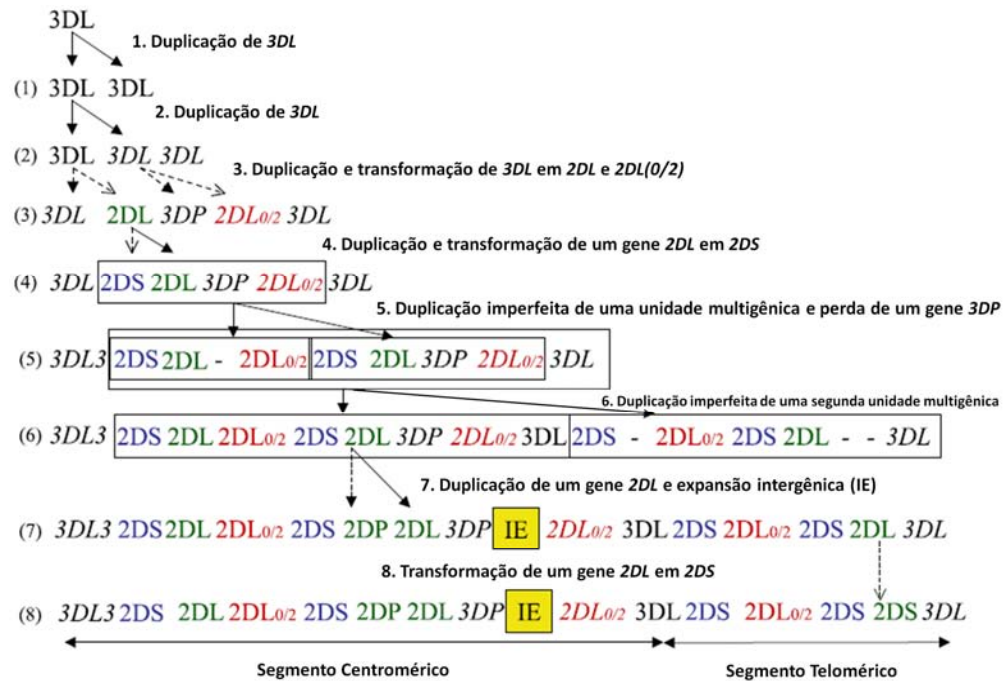


FIGURA 8. Esquema hipotético para o surgimento dos genes *KIR* a partir do gene *KIR3DL* ancestral. Repetidos eventos de duplicações parciais ou completas de genes deram origem aos genes *KIR* conhecidos atualmente (8). **Fonte:** MARTIN et al., 2004.

A presença de retroelementos comuns a todas as espécies analisadas, na região dos genes *KIR*, indica que estes genes derivam de um ancestral comum, originado antes da divergência entre estas espécies. Além disso, a presença de elementos *Alu* localizados em regiões similares de todos os genes *KIR* também sugere que estes genes têm uma origem única. A presença adicional de elementos *Alu* recentes (*S-Alu*) é um indicativo de frequentes eventos de duplicação, levando a diversificação dos genes *KIR* (WILSON et al., 2000; SAMBROOK et al., 2005).

Comparações entre diferentes espécies de primatas demonstram que os genes *KIR* são bastante variáveis entre as espécies e que apenas três dos genes *KIR* conhecidos (*KIR2DL4*, *KIR2DS4* e *KIR2DL5*) foram preservados durante a evolução dos homínídeos. Estas mesmas análises mostram que estes genes sofreram, de forma acelerada, extensas alterações, o que condiz com a sua ampla diversidade alélica e genotípica, características das diferentes espécies (KHAKOO et al., 2000; RAJALINGAM et al., 2001, SAMBROOK et al., 2005; MOESTA et al., 2009; MIDDLETON e GONZELES, 2009). Diferenças nas frequências dos genes *KIR* em populações geograficamente próximas também são evidências da rápida evolução destes genes (SINGLE et al., 2007).

A relativa proximidade dos genes *KIR* no cromossomo, bem como a grande semelhança nas sequências gênicas e intergênicas são responsáveis por permutas não homólogas que dão origem aos diferentes haplótipos com um número variado de genes *KIR*. A permuta desigual também é responsável pela formação de novos genes *KIR*, através da recombinação de segmentos de diferentes genes já existentes (GARDINER et al., 2001; HAO e NEI, 2005; PARHAM, 2008; NORMAN et al., 2009; MIDDLETON e GONZELES, 2009).

Os genes *KIR* possuem grande similaridade na sua sequência de nucleotídeos (80-90%), o que favoreceu a ocorrência eventos de recombinação não homóloga durante sua história evolutiva, levando a duplicações e deleções, e facilitando a rápida diversificação destes genes (SCHILLING et al., 1998; MARTIN et al., 2004; NORMAN et al., 2009; MIDDLETON e GONZELES, 2009). Análises filogenéticas das sequências que codificam os domínios protéicos dos diferentes receptores *KIR*, revelam que eventos de recombinação, como a mistura de exons (*exon shuffling*), foi bastante comum durante sua história evolutiva (HERSHBERGER et al., 2005). Mutações gênicas (ou pontuais) também desempenharam um importante papel na diversificação dos genes *KIR* (SCHILLING et al., 1998; GARDINER et al., 2001).

A grande similaridade na sequência de nucleotídeos dos diferentes genes *KIR* também é observada na região promotora destes genes. A existência de promotores comuns entre estes genes pode explicar por que expressão dos mesmos ocorre de forma aleatória em diferentes células de um mesmo indivíduo. Polimorfismos presentes nessas regiões podem favorecer a transcrição em uma ou outra direção, levando a transcrição de um ou outro gene, que são justapostos, porém codificados em direções opostas (WILSON et al., 2000). Polimorfismos na região reguladora destes genes também podem alterar o nível de transcrição dos mesmos, inclusive de seus diferentes alelos. Apesar de sua grande similaridade, os genes *KIR* e seus alelos codificam proteínas que diferem quanto a ligação às moléculas MHC, a expressão na superfície celular, sinalização intracelular e conformação da proteína, conferindo diferentes especificidades funcionais às células *NK* (STEWART, BERGEN e TROWSDALE, 2003; BASHIROVA et al., 2006; LI et al., 2008).

A diversidade alélica dos genes *KIR*, assim como de outros genes, é gerada por mutações de ponto, por eventos de conversão gênica e por recombinação homóloga, porém a contribuição relativa de cada um destes mecanismos varia de um gene para outro (THOMAS et al., 2008). Embora o polimorfismo alélico tenha sido descrito para todos os genes *KIR* ele parece ser muito mais intenso para os genes *KIR3DL1* e *KIR3DL2* (GARDNER et al., 2001; VILCHES e PARHAM, 2002).

O agrupamento dos diferentes genes *KIR* em haplótipos representa uma das principais fontes de diversidade destes genes dentro de uma mesma espécie. Repetidos eventos de recombinação são responsáveis pela formação de haplótipos constituídos por diferentes genes *KIR*, de modo que apenas indivíduos proximamente relacionados possuem genótipos *KIR* idênticos (KHAKOO et al., 2000; SHILLING et al., 2002; PARHAM, 2005b).

Aparentemente todos os genes *KIR*, com exceção de *KIR2DL4* apresentam uma sequência moderadamente repetitiva de 19-20 nucleotídeos, rica em conteúdo C+G, localizada geralmente no primeiro intron de cada gene (WILSON et al., 2000). Além disso, outras sequências repetitivas são bastante comuns na região dos genes *KIR*, como as sequências *Alus*, *SINEs* (Short Interspersed Nuclear Elements) e *LINEs* (Long Interspersed Nuclear Elements) (Figura 9). Em conjunto estas sequências repetitivas promovem instabilidades cromossômicas, levando a constantes reorganizações internas dos genes *KIR* e facilitando eventos de conversão gênica e recombinação não-homóloga (WILSON et al., 2000; TRAHERNE et al., 2010). A grande similaridade de sequência e a mesma orientação cromossômica dos diferentes genes *KIR* também favorecem a ocorrência de processos de conversão gênica entre estes genes (WILSON et al., 2000).

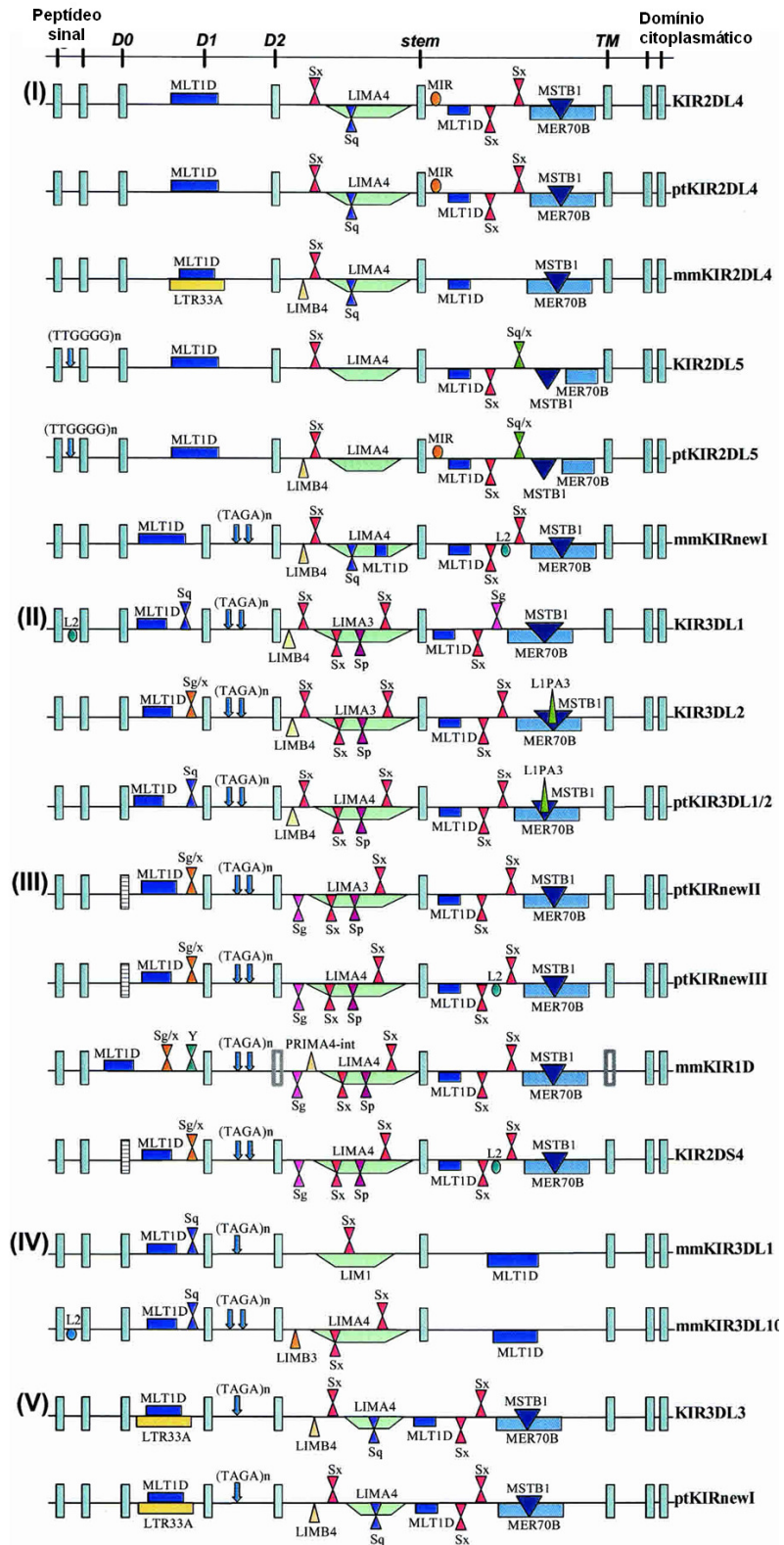


FIGURA 9. Representação esquemática da região intrônica dos genes *KIR*, mostrando a presença de retroelementos e sequências *ALU*. Os genes estão agrupados de acordo com as suas linhagens (I, II, III, IV e V). **Fonte:** SAMBROOK et al., 2005.

Durante a evolução dos genes *KIR*, a sua diversidade foi mantida nas diferentes espécies de primatas, principalmente pela ação de seleção positiva balanceadora, atuando sobre estes genes. (SALTER et al., 1997; HUGHES, 1999; HUGHES, 2002; SINGLE et al., 2007; GENDZECHADZE et al., 2009). Entretanto, eventos de seleção positiva direcional, seguidos por períodos de seleção negativa também podem ter ocorrido durante a sua história evolutiva, particularmente sobre os genes *KIR* ativadores. Os genes *KIR* inibidores apresentam uma diversidade significativamente superior, quando comparados aos genes *KIR* ativadores em todas as espécies de primatas. Essa diferença pode ser resultado da origem recente dos genes *KIR* ativadores, os quais podem ser formados por processos de recombinação entre genes *KIR* inibidores, de forma recorrente. Esses genes *KIR* ativadores estão sujeitos a seleção natural flutuante, especialmente seleção purificadora, que leva a diminuição das suas frequências e eventualmente sua perda (ABI-RACHED e PARHAM, 2005). Evidências dessa dinâmica de surgimento e desaparecimento dos genes *KIR* ativadores são fornecidas por estudos em populações humanas que mostram que a maioria dos genes *KIR* ativadores possuem frequências reduzidas quando comparados com os genes *KIR* inibidores (ABI-RACHED e PARHAM, 2005).

O papel central atribuído à interação KIR:HLA na modulação das respostas imunes sugere que estes locos evoluíram conjuntamente mantendo interações funcionais apropriadas, de modo que algumas moléculas codificadas por determinados alelos *KIR* possuem maior ou menor afinidade às diferentes moléculas HLA (KHAKOO et al., 2002; GENDZECHADZE et al., 2009). Alterações nas frequências das moléculas HLA de classe I, em resposta a pressões seletivas ou a fatores demográficos, podem afetar a evolução, através da manutenção ou expansão, do repertório dos genes *KIR* que interagem com estas moléculas HLA (HOLLANBACH et al., 2010). A análise das sequências dos diferentes genes *KIR* indica que as principais alterações ocorreram nos sítios de ligação às moléculas MHC, como resultado da ação de diferentes fatores evolutivos (HSU et al., 2002; MARTINEZ-BORRA e KHAKOO, 2008). Essas alterações influenciam a ligação KIR:HLA e podem estar sob ação de seleção natural, implicando na diversificação dos genes *KIR*.

O sistema *HLA* inclui os genes mais polimórficos do genoma humano e sua diversidade foi mantida nas populações por meio de seleção balanceadora, embora inúmeros eventos de seleção purificadora e de deriva genética também tenham ocorrido durante sua evolução. Hipóteses sugerem que eventos de seleção positiva sobre os genes *HLA*, favorecendo alelos vantajosos na defesa do organismo contra determinadas doenças, desempenharam um importante papel na evolução e diversificação dos genes *KIR* (DENIS

et al., 2005; GENDZEKHADZE et al., 2006; GENDZECHADZE et al., 2009; HOLLANBACH et al., 2010). A evolução dos genes *KIR* ocorreu, portanto, através da forte pressão seletiva sobre as moléculas MHC de classe I, e também, por meio de eventos de recombinação, conversão gênica e mutações de ponto (MARTINEZ-BORRA e KHAKOO, 2008).

A idéia de coevolução entre *HLA* e *KIR* é sustentada por estudos que associam genótipos destes genes com doenças, bem como por estudos de genética comparativa entre diferentes espécies de primatas. Vários trabalhos forneceram evidências sobre a existência de coevolução tanto alélica quanto genotípica e haplotípica entre *HLA* e *KIR* (HALFPENNY et al., 2004; CARR, PANDO e PARHAM, 2005; NORMAN et al., 2007; SINGLE et al., 2007; MARTIN et al., 2007; THANANSHAI et al., 2007; GENDZECHADZE et al., 2009; HOLLANBACH et al., 2010). Em alguns casos é evidente a ação da seleção direcional atuando a favor de determinados haplótipos, genótipos ou alelos (YAWATA et al., 2006; NORMAN et al., 2007; HOLLANBACH et al., 2010). Single e colaboradores (2007), analisando a diversidade global dos genes *KIR* e *HLA*, observaram uma correlação negativa entre os *KIR* ativadores e seus ligantes. Porém, na maioria das populações estudadas, as evidências sugerem que os genes *KIR* estão sob ação da seleção balanceadora e sua diversidade é mantida (TONEVA et al., 2001; PARHAM, 2005b; YAWATA et al., 2006; SINGLE et al., 2007; NORMAN et al., 2007; RAJALINGAM et al., 2008; HOLLANBACH et al., 2010).

As frequências, tanto alélicas como genotípicas e haplotípicas, dos genes *KIR* podem variar enormemente entre populações como resultado de diferentes pressões seletivas, bem como pela ação da deriva genética, do efeito fundador e do fluxo gênico (RAJALINGAM et al., 2008; KULKARNI et al., 2008). A variação nas frequências dos alelos para muitos dos genes *KIR*, entre diferentes populações, permite que esses genes sejam utilizados como marcadores de ancestralidade, evidenciando a história evolutiva da espécie humana (GARDNER, 2007; DENIS et al., 2005). Populações com histórias biológicas distintas podem mostrar diferenças nas frequências dos genes *KIR* devido a diferentes fatores evolutivos endógenos e exógenos aos quais estão submetidas.

A análise da diversidade global dos genes *KIR* revela diferenças entre populações com diferentes ancestralidades (TONEVA et al., 2001; YAWATA et al., 2006; GENDZEKHADZE et al., 2006; SINGLE et al., 2007; FLORES et al., 2007; RAJALINGAM et al., 2008; KULKARNI et al., 2008; HOLLANBACH et al., 2010). Estas diferenças são menos intensas entre populações geograficamente próximas, o que está diretamente relacionado com a história evolutiva das populações. Embora as frequências dos haplótipos *KIR* pertencentes aos haplogrupos A e B variem entre diferentes populações, como resultado,

provavelmente, da ação de diferentes pressões seletivas atuantes durante as migrações pré-históricas, de modo geral, a frequência de indivíduos heterozigotos para os haplótipos A e B é elevada em todas as populações humanas estudadas (HOLLANBACH et al., 2010).

2.1.1.3 *KIR3DL1* e *KIR3DS1*

Os genes *KIR3DL1* e *KIR3DS1* apresentam 97% de similaridade em sua sequência de nucleotídeos e aparentemente ocupam a mesma região genômica em diferentes haplótipos (WILSON et al., 2000). Estudos de segregação indicam que eles são alelos do mesmo loco (WILSON et al., 2000; GARDINER et al., 2001). Entretanto, já foram descritos haplótipos raros nos quais ocorre a presença ou ausência de ambos os genes (GARDINER et al., 2001; MIDDLETON e GONZELES, 2009). Em sul-asiáticos, haplótipos contendo ambos os genes (*KIR3DL1* e *KIR3DS1*) representam cerca de 10% dos haplótipos *KIR*, tendo surgido possivelmente por permuta desigual (CRUM et al., 2000; GARDINER et al., 2001; NORMAN et al., 2002).

KIR3DS1 surgiu provavelmente por duplicação seguida de deleção do gene *KIR3DL1* (GARDINER et al., 2001). Embora *KIR3DS1* apresente frequência elevada em todas as populações estudadas, sua diversidade é significativamente inferior que a observada para *KIR3DL1*. A ação de seleção purificadora sobre *KIR3DS1* pode explicar por que este gene é pouco variável, mesmo tendo surgido há milhares de anos. Neste trabalho, *KIR3DL1* e *KIR3DS1* são tratados como alelos do mesmo loco e serão denominados *KIR3DL1/S1*.

A evolução dos genes *KIR* em humanos, a partir da divergência entre os chimpanzés e o homem, ocorreu de forma bastante acelerada. O gene *KIR3DL1/S1* evoluiu a partir de três linhagens ancestrais (*KIR3DL1*005*, *KIR3DL1*015* e *KIR3DS1*013*), que permanecem bem representadas em todas as populações por mais de três milhões de anos (NORMAN et al., 2007; PARHAM, 2008; NORMAN et al., 2009). A ocorrência destas três linhagens em todas as populações estudadas é um claro indício de que estas linhagens evoluíram sob ação de seleção balanceadora na espécie humana. A população africana representa uma exceção e nesta população existem evidências de seleção direcional a favor da linhagem *KIR3DL1*015*, que apresenta uma frequência significativamente maior que as demais. Durante a evolução deste gene outras linhagens surgiram por recombinação entre as linhagens ancestrais (NORMAN et al., 2007; PARHAM, 2008; NORMAN et al., 2009).

A inibição das células *NK* é fundamental para a manutenção da tolerância aos constituintes próprios. Assim, os genes *KIR* que codificam receptores inibidores estão

presentes em virtualmente todos os indivíduos. Por outro lado, são conhecidos vários haplótipos que não possuem genes *KIR* ativadores (MIDDLETON e GONZELES, 2009). Dentre os genes que codificam receptores *KIR* inibidores, o gene *KIR3DL1* possui grande importância na modulação da ativação das células *NK* e sua frequência é elevada em todas as populações estudadas (SINGLE et al., 2007). O receptor ativador *KIR3DS1* também possui frequências elevadas e sua importância na ativação das células *NK* no combate a infecções já foi verificada em estudos anteriores (p. ex. MARTIN et al., 2007)

Para a maioria das populações investigadas neste trabalho as frequências do gene *KIR3DL1/S1* foram determinadas em estudos anteriores do nosso grupo (Augusto et al. manuscrito em preparação; dados não publicados) e são mostradas na tabela 1. Nestes mesmos estudos, através da comparação entre as frequências dos genes *KIR* e dos ligantes HLA, nenhuma correlação entre essas frequências foi encontrada. Embora não tenham sido encontradas evidências de coevolução entre os locos *KIR* e *HLA*, como indicativo de seleção natural, todos os indivíduos analisados possuem pelo menos uma combinação *KIR*:HLA funcional, ou seja, em todos os indivíduos, pelo menos um receptor *KIR* e seu respectivo ligante foram observados.

TABELA 1. Frequências dos genes *KIR3DL1* e *KIR3DS1* nas populações estudadas neste trabalho.

| | EURO-BR | ASI-BR | GKW | GND | GRC | KIV | KRC |
|----------------|---------|--------|-----|-----|-----|-----|-----|
| <i>KIR3DL1</i> | 93% | 23% | 81% | 86% | 69% | 78% | 80% |
| <i>KIR3DS1</i> | 35% | 43% | 71% | 56% | 82% | 56% | 66% |

EURO-BR - Eurobrasileiros; ASI-BR - Brasileiros descendentes de Asiáticos; GKW - Guarani Kaiowá; GND - Guarani Nandeva; GRC - Guarani M'byá; KIV - Kaingang de Ivaí; KRC - Kaingang de Rio das Cobras. **Fonte:** Augusto et al. manuscrito em preparação; dados não publicados.

KIR3DL1/S1 é um gene bastante polimórfico cujo produto se liga a moléculas HLA com epítipo Bw4, que representam cerca de 33% das moléculas HLA-B e 20% das moléculas HLA-A. O epítipo Bw4 é definido pelos aminoácidos presentes nas posições 77, 80, 81, 82 e 83 da cadeia polipeptídica. Este epítipo pode apresentar um resíduo de isoleucina ou treonina na posição 80, conferindo diferentes afinidades de interação aos seus receptores (WAN et al., 1986). Moléculas com isoleucina nesta posição têm, geralmente, uma maior afinidade aos seus receptores se comparadas com as moléculas que apresentam treonina (KULKARNI, MARTIN e CARRINGTON, 2008). Os alelos *KIR3DS1*

parecem ligar-se exclusivamente a moléculas HLA-Bw4 com isoleucina na posição 80 (MARTIN et al., 2007; ALTER et al., 2007), embora estudos mais recentes demonstrem que não haja interação entre ligantes HLA-Bw4 e o receptor KIR3DS1 (MORVAN et al., 2009).

As frequências dos genes *HLA* e *KIR* variam de uma população para outra e a importância da interação KIR:HLA na modulação da atividade das células *NK* sugere que estes locos podem ter evoluído conjuntamente. Vários estudos evidenciaram a existência de coevolução entre alelos de *KIR3DL1/S1* e *HLA - A/B* com epítipo *Bw4* (PANDO e PARHAM, 2005; NORMAN et al., 2007; MARTIN et al., 2007; THANANSHAI et al., 2007; CARR, THOMAS et al., 2008). A expressão dos alelos *KIR3DL1/S1* na superfície das células *NK* é determinada por polimorfismos presentes ao longo da sequência de nucleotídeos e por diferenças na região promotora dos diferentes alelos.

São conhecidos 84 alelos para o gene *KIR3DL1/S1* (<http://www.ebi.ac.uk/ipd/kir/>), cujas frequências variam entre as populações estudadas. O polimorfismo deste gene tem uma importância funcional significativa na regulação das células *NK* (YAWATA et al., 2006). Variações na região transmembrana e nos três domínios extracelulares, principalmente, estão relacionadas com o nível de expressão dos diferentes alelos *KIR3DL1/S1* na superfície celular e com a afinidade pelo ligante HLA. Dessa forma, os diferentes alelos *KIR3DL1* possuem um potencial inibidor distinto.

Os polimorfismos na região transmembrana e nos domínios extracelulares podem alterar a conformação da proteína causando a sua retenção no interior das células, ou impedir o seu ancoramento na membrana celular, de modo que alguns alelos são superexpressos (por exemplo, *KIR3DL1*002*), enquanto outros são expressos em baixos níveis (por exemplo, *KIR3DL1*005*) ou mesmo não expressos (por exemplo, *KIR3DL1*004*). Polimorfismos na região promotora do gene *KIR3DL1/S1*, além de mecanismos epigenéticos como a metilação, também podem estar relacionados com o nível de expressão dos diferentes alelos na superfície das células (GARDINER et al., 2001; CHAN et al., 2003; STEWART, BERGEN e TROWSDALE, 2003; LI et al., 2008; GATTAZZO et al., 2010). Além disso, as variações nos domínios extracelulares podem alterar a afinidade da interação com HLA (GARDINER et al., 2001; PANDO et al., 2003; O'CONOR et al., 2007).

A diversidade alélica do gene *KIR3DL1/S1* é resultado, principalmente, de eventos de recombinação e de mutações de ponto que ocorreram durante a sua história evolutiva (SCHILLING et al., 1998; GARDINER et al., 2001; NORMAN et al., 2009). As mutações de ponto, embora sejam importantes para a diversificação deste gene, contribuíram em menor nível para a sua diversidade alélica do que a recombinação. Fatores evolutivos como a

seleção e a deriva também são importantes para explicar a diversidade do gene *KIR3DL1/S1* (GARDINER et al., 2001). Comparações de sequência de aminoácidos entre diferentes indivíduos revelam a existência de 21 posições, distribuídas entre os três domínios extracelulares (D0, D1 e D2), que estão sob seleção positiva, sendo que algumas destas posições são variáveis, enquanto outras são conservadas de um indivíduo para outro (Figura 10). Muitas destas posições que possivelmente estão sob seleção positiva correspondem a resíduos que participam da interação com as moléculas HLA (NORMAN et al., 2007).

FIGURA 10. Sequência de aminoácidos da proteína codificada pelo alelo *KIR3DL1*015* mostrando a localização das posições que estão sob ação de seleção positiva nos três domínios extracelulares (D0, D1 e D2) dos receptores KIR3DL1/S1. Estas posições estão indicadas pelos símbolos \dagger e \ddagger e os retângulos representam os sítios de contato com as moléculas HLA. **Fonte:** NORMAN et al., 2007.

Embora o domínio zero (D0) destes receptores não participe diretamente da interação KIR:HLA ele atua na modulação da interação D1/D2 (domínios 1 e 2) com o ligante. Variações em D0, principalmente nas posições 49-52 da sequência de aminoácidos, podem alterar a conformação dos domínios D1 e D2 influenciando a ligação com as moléculas HLA-Bw4 (KHAKOO et al., 2002; PANDO et al., 2003; NORMAN et al., 2007).

A diversidade funcional dos alelos do gene *KIR3DL1/S1*, que é determinada principalmente pelos três domínios extracelulares, evoluiu sob ação de seleção balanceadora (HOLLANBACH et al., 2010). Porém, na população africana, ao contrário das demais, os domínios D1 e D2 aparentemente estão sob ação de seleção direcional a favor da linhagem *3DL1*015* e apenas para D0 existem evidências de seleção balanceadora (NORMAN et al., 2007).

As frequências alélicas de *KIR3DL1/S1* diferem de uma população para outra, sendo que um alelo que possui frequência elevada em determinada população pode estar em baixa frequência em outra população. Essas diferenças interpopulacionais podem ocorrer devido à seleção natural, como resultado da interação com as moléculas HLA, podendo, cada combinação KIR:HLA ter adquirido diferentes padrões de interação e ter sido diferentemente selecionadas entre as populações. Além disso, alguns alelos são exclusivos de determinadas populações. Gendzekhadze e colaboradores (2006) investigando a diversidade dos genes *KIR* em populações ameríndias da Venezuela identificaram a presença de cinco alelos *KIR3DL1* e dois alelos *KIR3DS1*. Estes alelos se distribuem uniformemente entre os haplótipos identificados nestas populações. Neste estudo, uma nova variante *KIR3DS1* foi identificada (*3DS1*047*). Essa variante é caracterizada por uma substituição não sinônima de nucleotídeo no domínio extracelular D1 e parece ser exclusiva dos indígenas da tribo Warao.

Apesar de em alguns estudos a diversidade alélica de *KIR3DL1/S1* em populações humanas já ter sido investigada (por exemplo, GENDZEKHADZE et al., 2006; HALFPENNY et al., 2004; FLORES et al., 2007), a diversidade deste gene, bem como a implicação funcional desta diversidade, permanece desconhecida para a grande maioria das populações.

2.2 POPULAÇÕES HUMANAS

A espécie humana surgiu na África há aproximadamente 200 mil anos e, há apenas 100 mil anos passou a colonizar outros continentes, ocupando a Europa, Ásia e posteriormente a América (TISHKOFF et al., 2009). Durante as migrações humanas para fora da África, as populações sofreram efeito de deriva genética devido à redução do tamanho efetivo da população (efeito gargalo de garrafa), com consequente diminuição da sua diversidade genética (KIDD et al., 1998; LONJOU et al., 2003; CAVALLI-SFORZA e FELDMAN, 2003). Dessa forma, individualmente as populações apresentam, em geral, apenas uma parte da diversidade genética encontrada dentro do continente africano.

2.2.1 POPULAÇÕES AMERÍNDIAS

Ao contrário das populações urbanas, nas quais o fluxo gênico parece ter um papel preponderante, nas populações indígenas, que são populações pequenas e geneticamente isoladas, a ação do efeito fundador e da deriva genética desempenha um importante papel na determinação da sua diversidade genética. Estudos genéticos de populações ameríndias, aliados a abordagens arqueológicas, paleontológicas e linguísticas levaram à elaboração de hipóteses acerca do povoamento das Américas.

O continente americano foi um dos últimos locais a ser povoado pela espécie humana, e a hipótese mais aceita entre os cientistas é de que o homem veio da Sibéria para a América, durante a última glaciação, através do estreito de Bering. Análises do DNA mitocondrial, cromossomo Y e outras regiões polimórficas apóiam mais de uma hipótese, e ainda hoje não há um consenso sobre a data da primeira migração, número de migrações e das rotas utilizadas.

Os resultados obtidos através do sequenciamento do genoma mitocondrial humano, em nativos americanos, sugerem que o povoamento do continente ocorreu por uma única onda migratória originada provavelmente no nordeste asiático há cerca de 20 mil anos. Antes da entrada na América, ocorreu um confinamento por cerca de cinco mil anos na Beríngia, evidenciado por um pequeno efeito de gargalo de garrafa. A entrada no continente americano ocorreu há aproximadamente 18-19 mil anos, quando houve um grande aumento populacional (FAGUNDES, KANITZ e BONATTO, 2008; FAGUNDES et al., 2008).

Dados obtidos a partir da análise da morfologia de crânios indicam que pelo menos duas populações biologicamente distintas colonizaram as Américas durante a transição do Pleistoceno para o Holoceno (há cerca de 12 mil anos). A primeira teria a morfologia do crânio similar à encontrada hoje na África e Austrália, e a segunda teria a morfologia similar à encontrada no nordeste asiático (NEVES e HUBBE, 2005; NEVES, HUBBE e PILÓ, 2007). Entretanto, em uma revisão que uniu dados de morfologia e de genética, foi sugerido que a população que colonizou a América apresentava grande diversidade craniofacial, e as morfologias craniofaciais paleoamericanas e mongolóides são, na realidade, extremos de uma variação morfológica contínua (GONZÁLEZ-JOSÉ et al., 2008).

Após a entrada no continente americano, as migrações ocorreram do norte em direção ao sul, mas não há consenso a respeito das rotas utilizadas. Evidências genéticas e de morfologia craniofacial indicam a existência de duas rotas costeiras, uma pelo oceano Pacífico e outra pelo oceano Atlântico (PUCCIARELLI et al., 2008).

No Brasil, estima-se que havia entre dois a quatro milhões de indígenas, distribuídos por mais de 1000 povos na época da chegada dos europeus, mas hoje restam apenas 230 povos, totalizando cerca de 600 mil indivíduos (0,2% da população brasileira), falantes de mais de 180 línguas distintas. Dentre esses, 51 mil são Guarani (FUNAI, 2008) e 33 mil são Kaingang (Funasa, 2009). Os Guarani e Kaingang são as duas principais tribos indígenas do sul do Brasil. Possuem língua e cultura distintas, e mesmo quando habitam a mesma reserva indígena, moram em aldeias diferentes e tendem a permanecer endogâmicos. No Paraná existem aproximadamente nove mil indígenas distribuídos por 17 áreas, incluindo as etnias Kaingang, Guarani (Ñandeva e M'byá) e Xetá. No Mato Grosso do Sul vivem nove etnias indígenas, dentre as quais Guarani Kaiowá e Ñandeva, totalizando aproximadamente 55 mil indivíduos (INSTITUTO SOCIOAMBIENTAL).

Em vários estudos foi analisado o polimorfismo dos genes *HLA*, outros genes do *MHC* e marcadores microssatélites desta região, em populações ameríndias. Os resultados mostraram que a diversidade dos locos analisados em populações indígenas é menor do que o de outras populações, e o desequilíbrio de ligação entre os locos é maior (WANG et al., 2007; OLIVEIRA et al., 2008). Dentre as populações ameríndias foram encontrados alelos *HLA* exclusivos destas populações, que surgiram a partir dos alelos fundadores principalmente por eventos de conversão gênica (BELICH et al., 1992; PARHAM et al., 1997).

Por se tratarem de genes importantes na regulação da atividade das células *NK*, os genes *KIR* podem estar sob ação de diversos fatores evolutivos, sendo especialmente

interessantes para os estudos de genética de populações. Alguns estudos que investigaram os haplótipos formados pelos diferentes genes *KIR* foram desenvolvidos em populações ameríndias, e, em comparação com as populações urbanas, as populações ameríndias apresentam uma diversidade genética (combinações dos diferentes genes *KIR*) significativamente menor (GUTIERREZ-RODRIGUES et al., 2006; GENDZEKHADZE et al., 2006; FLORES et al., 2007). Estes mesmos estudos revelam que alguns haplótipos raros nas populações urbanas podem ter uma frequência consideravelmente maior nas populações indígenas. Os efeitos demográficos parecem ser os principais agentes que contribuem para a ampla diversidade haplotípica dos genes *KIR* nas populações urbanas (FLORES et al., 2007).

Nas populações ameríndias, nas quais o fluxo gênico é muito baixo, ou mesmo ausente, a pressão de seleção poderia explicar porque a diversidade haplotípica de *KIR* é relativamente baixa. Durante a história destas populações podem ter sido selecionados determinados haplótipos em detrimento de outros. Por outro lado, fatores estocásticos como o efeito fundador e a deriva genética podem ter contribuído enormemente para a diversidade observada nestas populações, que, geralmente, possuem tamanho reduzido e são geograficamente isoladas.

Considerando a importância na regulação da atividade do sistema imune inato e na reprodução, a combinação *KIR*:HLA deve ter representado um importante papel na sobrevivência e manutenção das populações ameríndias, estando, estes locos, sob ação de diferentes fatores evolutivos (YAWATA et al., 2006; GENDZEKHADZE et al., 2006).

2.2.1.1 As populações indígenas Kaingang e Guarani

Até o contato com os colonizadores, os Kaingang ocupavam grandes áreas no sul do Brasil (sempre na região central dos estados) e na Argentina. Viviam principalmente da agricultura, além da caça, pesca e coleta. Eram seminômades e guerreiros. Hoje estão restritos a 30 terras indígenas, nos estados de Paraná, Rio Grande do Sul, Santa Catarina e São Paulo, muitas vezes compartilhadas com outras tribos, e são a terceira etnia indígena quanto ao tamanho da população. A tribo Kaingang é dividida em dois grupos, Kamé e Kairu, sendo que os indivíduos de um grupo devem se casar com os do grupo oposto, e os descendentes recebem a filiação paterna, ou seja, as famílias são patrilineares. Apesar do padrão patrilinear de descendência, a forma tradicional de residência entre os Kaingang é a matrilocalidade - após o casamento o genro passa a morar na casa do sogro. Sua língua é

o Kaingang, classificado na família Jê, tronco linguístico Macro-Jê (INSTITUTO SOCIOAMBIENTAL) (Figura 11).

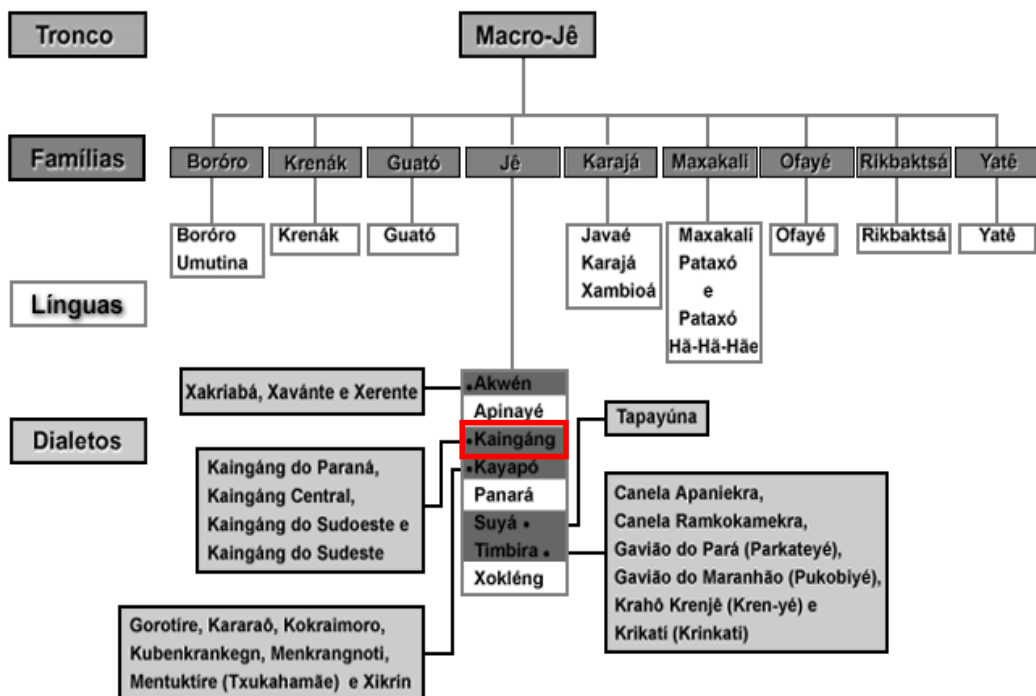


FIGURA 11. Famílias, línguas e dialetos pertencentes ao tronco linguístico MACRO-JÊ. Em destaque os dialetos Kaingang das populações analisadas neste estudo. **Fonte:** INSTITUTO SOCIOAMBIENTAL.

Neste estudo foram analisados Kaingang residentes em duas áreas de reserva indígenas: Rio das Cobras (25°18'S, 52°32'O), a maior do Paraná, com 18.681,98 ha, localizada no município de Nova Laranjeiras, que abriga aproximadamente de 2.397 indivíduos, dos quais cerca de um terço são Kaingang e os demais são Guarani e Xetá; e na área indígena Ivaí (24°30'S, 51°40'O), localizada no município de Manoel Ribas, com 7.306,34 ha, que abriga cerca de 1.168 pessoas.

Os Guarani eram nômades antes da colonização, viviam principalmente da agricultura e habitavam uma grande área na região litorânea dos estados do Rio Grande do Sul, Santa Catarina, Paraná, São Paulo e também no estado do Mato Grosso do Sul, além de países vizinhos. Os Guarani têm como base de sua organização social, econômica e

política, a família extensa, isto é, grupos macro familiares que detêm formas de organização da ocupação espacial dentro dos *tekoha* (território habitado) determinadas por relações de afinidade e consanguinidade. A família é composta pelo casal, filhos, genros, netos, irmãos e constitui uma unidade de produção e consumo. Os casamentos devem ocorrer entre membros de famílias distintas, havendo regras explícitas de proibição de casamentos dentro de uma mesma família. Hoje vivem no sul e sudeste do Brasil, Paraguai, Argentina e Bolívia, e são a maior etnia indígena brasileira. Sua língua é o Guaraní, da família Tupi-Guarani, tronco linguístico Tupi (Figura 12) (INSTITUTO SOCIOAMBIENTAL).

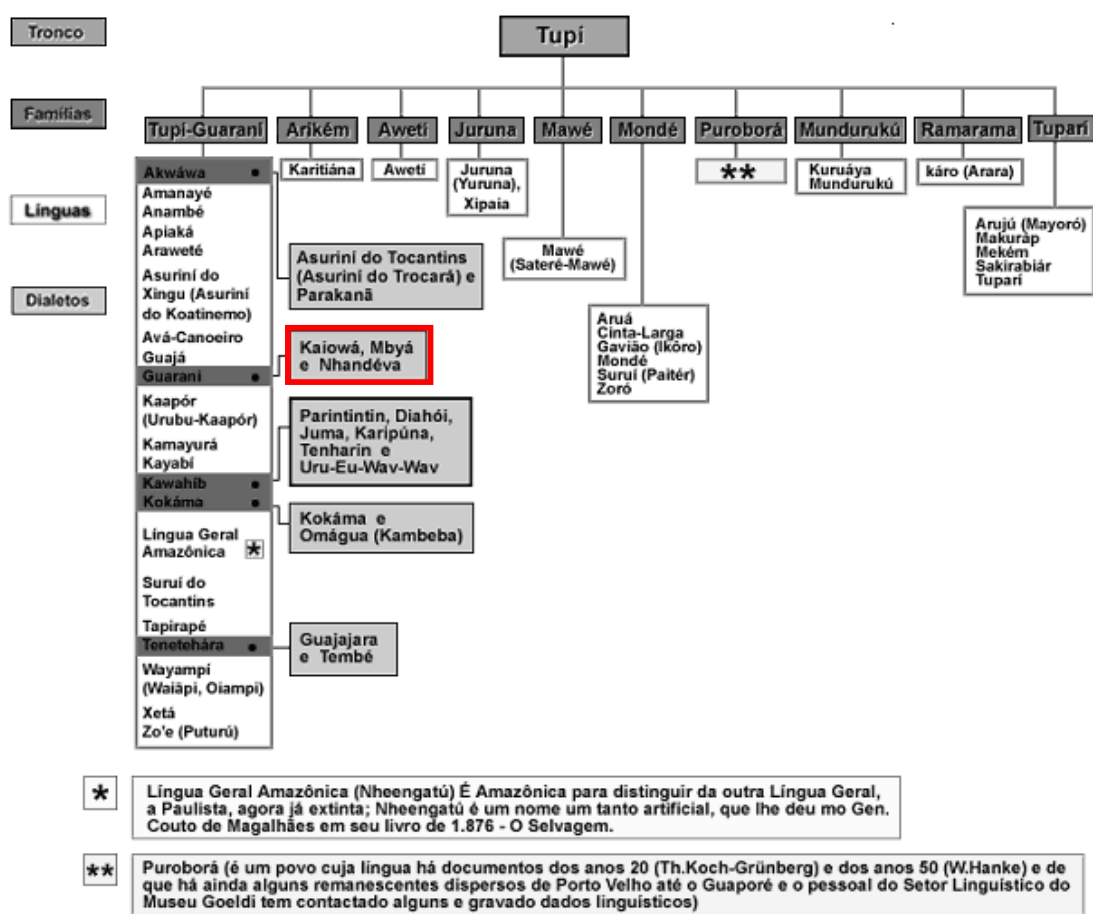


FIGURA 12. Famílias, línguas e dialetos pertencentes ao tronco linguístico TUPI. Em destaque os dialetos Guaraní das populações analisadas neste estudo: Guaraní M'byá, Kaíowá e Nhandéva. **Fonte:** INSTITUTO SOCIOAMBIENTAL.

Os Guaraní analisados neste estudo pertencem a três grupos distintos, nomeados pelo dialeto que falam: Guaraní M'byá (da área indígena Rio das Cobras, município de Nova

Laranjeiras, PR); Guarani Kaiowá, das aldeias Amambai (23°06'S; 55°12'O) e Limão Verde (23°12'S; 55°06'O), que abrigam cerca de 6.663 e 669 indivíduos, respectivamente, localizadas no município de Amambai, Mato Grosso de Sul; e Guarani Nandeva, de Amambai e Porto Lindo (23°48'S; 54°30'O), MS, que abriga aproximadamente 4.800 pessoas.

Estudos realizados com as tribos Kaingang e Guarani mostraram que, além da diferença entre populações indígenas e urbanas, elas são geneticamente diferentes entre si, apesar de atualmente conviverem em muitas áreas de reserva indígena (PETZL-ERLER et al., 1993; WEG-REMERS et al., 1997). Guarani e Kaingang também diferem quanto às frequências de linhagens de DNAm e em menor grau quanto a haplogrupos de cromossomo Y. A quantidade de variação interpopulacional também difere entre as duas tribos, resultado compatível com dados linguísticos e históricos, que indicam que a migração do grupo Tupi foi mais antiga do que a do grupo Jê (MARRERO et al., 2007).

O fluxo gênico entre populações indígenas e nãoindígenas é muito baixo, e a taxa de miscigenação estimada através da análise de genes *HLA* de classe II foi de 7% para Kaingang, 4% para Guarani M'byá, 14% para Guarani Nandeva e ausente para Guarani Kaiowá (TSUNETO et al., 2003).

2.2.2 POPULAÇÃO BRASILEIRA

A população brasileira é heterogênea e miscigenada, diferindo de uma região para outra. É composta por indivíduos de ascendência européia, africana subsaariana, ameríndia e oriental, em diferentes proporções dependendo da região estudada. Na região sul, de onde provém a amostra analisada neste estudo, há predominância do componente europeu. Segundo o Instituto Brasileiro de Geografia e Estatística (IBGE, 2007), a população da região sul do Brasil é composta por 78,7% de brancos, 4,3% de pretos, 16,3% de pardos e 0,7% de amarelos e indígenas. Porém a auto-classificação dos indivíduos em brancos, mulatos e negros não corresponde às mesmas proporções de ancestralidade européia e africana em diferentes regiões do país, mostrando que a cor da pele é um marcador pouco informativo da ancestralidade genômica de indivíduos (PARRA et al., 2002; PIMENTA et al., 2006).

A população da região metropolitana de Curitiba (com aproximadamente 3 milhões e 300 mil habitantes, conforme dados coletado pelo IBGE em 2009) é composta por 84,65%

de brancos, 11,34% de pardos, 2,49% de negros e 1,07% de amarelos. Para o percentual restante não foi feita declaração (Prefeitura Municipal de Curitiba). Entretanto, estudos genéticos demonstram que, de forma geral, a cor da pele é um marcador pouco informativo acerca da ancestralidade genômica das populações humanas. Para população de Curitiba e região metropolitana, um estudo, analisando genes *HLA* de classe I (*HLA-A*, *-B* e *-C*) e II (*HLA-DR* e *-DQ*), mostrou que o componente europeu é predominante entre os brancos (80,6%), sendo que a contribuição de africanos é de 12,5% e a de ameríndios é de 7,0%. Já para os mulatos o componente predominante foi o africano (49,5%), sendo as contribuições européia de 41,8% e ameríndia de 8,7% (PROBST et al., 2000). BRAUN-PRADO e colaboradores (2000), analisando genes *HLA* de classe I, mostraram que não só para os brancos, mas também para os mulatos, o componente europeu é predominante. A contribuição de europeus, africanos e ameríndios foi, respectivamente, 94%, 3% e 3% para brancos e 57%, 39% e 4% para mulatos.

3 OBJETIVOS E JUSTIFICATIVAS

A interação entre *KIR3DL1/S1* e as moléculas HLA-A e HLA-B é importante para a regulação da atividade das células *NK*, que têm grande importância nas respostas imunológicas. Além disso, poucos estudos investigaram as implicações funcionais da diversidade alélica do gene *KIR3DL1/S1* em populações ameríndias e em outras populações. Dessa forma, passa a ser significativa a contribuição de estudos populacionais caracterizando a variabilidade alélica dos genes *KIR* e suas implicações na imunidade e doenças. Além disso, o estudo da sequência nucleotídica dos genes é bastante útil para a compreensão da origem e dinâmica dos seus polimorfismos, auxiliando no entendimento dos mecanismos evolutivos que atuam sobre estes genes.

O objetivo deste trabalho é caracterizar a diversidade alélica do gene *KIR3DL1/S1* em populações ameríndias dos estados do Mato Grosso do Sul e do Paraná, bem como na população mestiça brasileira (de origem predominantemente européia e africana), na população de ancestralidade oriental e na população africana. Os objetivos específicos compreendem:

Determinar as frequências alélicas e genotípicas do gene *KIR3DL1/S1* nas populações estudadas.

Verificar se há evidências de coevolução entre as variantes encontradas e alelos *HLA* nestas populações, através da correlação de frequências.

Levantar hipóteses acerca da existência e origem de novos alelos do loco *KIR3DL1/S1*.

Comparar as frequências encontradas nas diferentes populações e estimar a distância genética entre elas, incluindo nesta análise outras populações para as quais já existam dados publicados.

Buscar, através da análise das sequências gênicas e das frequências alélicas, evidências indicativas de fatores evolutivos atuando sobre a diversificação do gene *KIR3DL1/S1*.

4 CAPÍTULO I - ARTIGO

High Polymorphism of *KIR3DL1/S1* gene in Amerindians and Evidences of Natural Selection in Functional Domains

Ricardo Dalla-Costa¹, Danillo Gardenal Augusto¹, Sheyla Mayumi Kuniwake¹, Liana Alves de Oliveira¹, Roseli Wassen¹, Angelica Beate Winter Boldt², Luiza Tamie Tsuneto³; Mary Carrington⁴ and Maria Luiza Petzl-Erler¹

¹ Department of Genetics, Federal University of Paraná, Curitiba, Brazil

² Immunopathology Laboratory, Federal University of Paraná, Curitiba, Brazil

³ Department of Clinic Analyses, University of Maringá, Paraná, Brazil

⁴ National Cancer Institute at Frederick, Frederick, United States of America

Keywords: *KIR*; genetic polymorphism; Amerindians, population genetics

Correspondence to: Maria-Luiza Petzl-Erler
Department of Genetics
Federal University of Paraná
Caixa Postal 19071
81531-980 Curitiba, Brazil
Tel. (+55 41) 33611746
Fax (+55 41) 33611793
e-mail: perler@ufpr.br

Abstract

The killer-cell immunoglobulin-like receptor (KIR) family is an important group of NK cell receptors that are expressed in a variegated manner on the NK cell population of a given individual. The interaction between KIR and HLA ligands can result in NK cell inhibition or activation and plays an important role in control of immune responses. The *KIR* genes are among the most polymorphic genes of the human genome and they have evolved rapidly, in a species-specific manner. At present, 84 alleles are assigned to the *KIR3DL1/S1* locus, and its products interact with the Bw4 motif of HLA-B and -A molecules. We analyzed 451 individuals, including Brazilians of predominantly European, of mixed but predominantly African and European, and of Asian ancestry, as well as five Amerindian populations from the Guarani and Kaingang groups, and an African population sample. Twenty five alleles were identified over all populations, including four previously unknown alleles. Despite the lowest genetic diversity in Amerindians, these populations presented high number of alleles and a considerable genetic diversity for the *KIR3DL1/S1* gene. Among the Amerindian populations analyzed, Guarani Ñandeva had the highest gene flow from Europeans and Africans and accordingly had the highest genetic diversity. Instead, the Guarani M'byá had the lowest diversity. All populations differed greatly for allele frequencies. *KIR3DS1*01301* was the most common allele in Amerindians, accounting for up to 53% of the Guarani M'byá alleles. Conversely, the lineage *3DS1* was rare in Africans, as reported by others. Molecular analyzes showed an excess of non-synonymous substitutions in exons 4 and 5, at sites predicted to interact with HLA molecules but not in the other regions analyzed. In conclusion, these results showed that *KIR3DL1/S1* is highly polymorphic in all populations analyzed and there are evidences of coevolution between *KIR* and *HLA* loci.

Introduction

Natural killer (NK) cells are a subset of lymphocytes that play a critical role in the innate immune response and also are important in homeostasis maintenance (Trinchieri, 1995). The activation or inhibition of NK cell functions depend on the balance between positive and negative signals from a diverse repertoire of membrane receptors (reviewed in Di Santo, 2006). Many of the NK cell receptors recognize MHC class I molecules (named also HLA class I in humans) that are expressed by most of the cells of an organism. The killer-cell immunoglobulin-like receptor (KIR) family is an important group of NK cell receptors that are expressed in a variegated manner on the NK cell population of a given individual (Chan et al., 2003). The HLA (Human Leukocyte Antigen) genotypes of an

individual also appear to be important for determining which KIRs will be expressed (Parham, 2006; Gardner, 2007).

The *KIR* genes are clustered in tandem on chromosome 19 (19q13.4) as haplotypes that usually are divided in two groups, named A and B (Ashworth et al., 1995; Trowsdale, 2001). Each of these haplogroups encompasses many haplotypes that differ not only by gene content but also by the alleles of the several loci. Individual haplotypes may be shared between populations, revealing a biological link between human groups, while others are characteristic of one or a few populations, reflecting their local origin (<http://www.allelefrequencies.net/>). Haplotypic differences among populations reflect the possible roles of natural selection and of demographic factors resulting in random genetic drift and the founder effect. (Toneva et al., 2001; Shilling et al., 2002; Denis et al., 2005; Gendzekhadze et al., 2006; Flores et al., 2007; Middleton et al., 2007; Rajalingam et al., 2008).

KIR gene content and allelic polymorphisms of the *KIR* and *HLA* genes could affect the functions and the specificity and sensibility of NK cells (Carr et al., 2005; O'Connor et al., 2007; Kim et al., 2008). Since both the *KIR* and the classical *HLA* genes are highly polymorphic and segregate independently, it is possible that certain combinations of *KIR* and *HLA* alleles could result in beneficial or deleterious interaction. In fact, many studies revealed the coevolution of *KIR* and *HLA* in world-wide populations (Hiby et al., 2004; Martin et al., 2007; Norman et al., 2007; Single et al., 2007; Thananchai et al., 2007; Yu et al., 2007; Kulkarni et al., 2008). *KIR*, as well as *HLA* genetic diversity, is most likely maintained by balancing or fluctuating natural selection, which acted diversifying the residues of KIR receptors that are involved in the contact with HLA molecules (Hugles, 2002; Hao et al., 2005; Norman et al., 2007).

Although *KIR* genes act epistatically, some of them appear to be more important for modulating NK cell function than others, because their high frequency in most investigated populations and their high level of allelic polymorphism. The *KIR3DL1* and *KIR3DS1* genes were assigned as alleles of the same locus (Wilson et al., 2000; Gardiner et al., 2001) that presents the highest number of alleles among all the *KIR* genes. The KIR3DL1/S1 receptors recognize HLA-A and HLA-B molecules with the Bw4 motif. The expression level of *KIR3DL1/S1* differs between alleles, which may present high, low or no expression on the cell surface. This variation may have an impact on the control of NK cell activity (Gardiner et al., 2001; Pando et al., 2003; Carr et al., 2005; Yawata et al., 2006; O'Connor et al., 2006). All KIR genes present nucleotide sequence variations distributed through all the length of the

gene. Polymorphisms observed in the cytoplasmatic tail (CYT), transmembrane region (TM) and extracellular immunoglobulin-like (Ig-like) domains (D0, D1 and D2) of the *KIR3DL1/S1* gene were correlated with the levels of expression at the NK cell surface (Gardiner et al., 2001; Pando et al., 2003; Thomas et al., 2008). Furthermore, variations on extracellular domains could alter the affinity of interaction with HLA ligands with functional consequences, being target for natural selection (Gardiner et al., 2001; Pando et al., 2003; O'Connor et al., 2007; Norman et al., 2007). Polymorphisms in the promoter region of *KIR3DL1/S1* alleles are also implicated on their expression (Li et al., 2008)

It has been suggested that alleles of the *KIR3DL1/S1* human gene originated mainly by recombination and secondarily by point mutations (Schilling et al., 1998; Gardiner et al., 2001). Although most of the *KIR3DL1/S1* alleles are shared between human populations, some have been described as characteristic of certain groups (e.g. Gendzekhadze et al., 2006). Since populations differ in the frequencies of *KIR3DL1/S1* alleles, the variants of this gene can be used as ancestry markers to investigate the relationship between different human groups.

The American continent was the last occupied by humans, about 20,000 years ago (Fagundes et al., 2008; Fagundes, Kanitz and Bonatto, 2008). Still in present days, many native Americans are grouped in small populations in which genetic drift and related factors, such as the founder and bottleneck effects, play major roles in their evolutionary history, in opposition to urban mixed populations, in which gene flow is a very important evolutionary factor. Regardless of some studies that have investigated allelic diversity of *KIR3DL1/S1* gene in populations (e.g., Gendzekhadze et al., 2006; Flores et al., 2007; Norman et al., 2007), its diversity remains unknown in most populations. In addition, little is known about the implications of their variation. Thus, based on the full length of the *KIR3DL1/S1* coding sequence, this study focuses on allelic diversity in populations of different ancestries, including Amerindians, Euro-Brazilians, Asians, and Africans, shedding light on the evolution of this gene's polymorphism in human populations.

Methods

Population samples

We analyzed 451 individuals from different ancestries. The Kaingang population sample is from Ivaí (n=60) (24°30'S, 51°40'W) and Rio das Cobras (n=72) (25°18'S,

52°32'W), both in Paraná state, Brazil. The Guarani M'byá (n=67) are also from Rio das Cobras. This reservation area is shared by Guarani M'byá and Kaingang. The Guarani Ñandeva (n=58) are from Amambaí (23°06'S; 55°12'W) and Porto Lindo (23°48'S; 54°30'W), in the state of Mato Grosso do Sul, Brazil. The Guarani Kaiowá (n=64) are from Limão Verde (23°12'S; 55°06'W) and also from Amambaí (Figure 1). A sample of 51 individuals with predominantly European and African ancestry, contacted at Curitiba, Paraná state, Brazil, was also analyzed. The population of Asian ancestry (n=56) also lives in Curitiba. The African population (n=23) forms a heterogeneous group, with individuals from various origins. Blood collection and DNA extraction were as described before (Petzl-Erler, Luz & Sotomaior, 1993; Petzl-Erler and McDevitt, 1994; Tsuneto et al., 2003).

Primers for the amplification by PCR and for DNA sequencing

Individuals analyzed were previously typed for presence and absence of *KIR* genes (unpublished data). The PCRs were performed to amplify the exons 1-5, 7 and 9 of the *KIR3DL1/S1* gene, using the following primer pairs: KIR3DL1/S1Ex1e2F (5'-cctgggagtgtaaatcatttg-3') and KIR3DL1/S1Ex1e2R (5'-ctgtgaccatgatcaccac-3'); KIR3DL1/S1Ex3F (5'-ttggccagagggccggtc-3') and KIR3DL1/S1Ex3R (5'-cctgtgaccatgatcaccay-3'); KIR3DL1/S1Ex4F (5'-tcccctggtgaaatcaggaga-3') and KIR3DL1/S1Ex4R (5'-gtgaccacgatgtccagggga-3'); KIR3DL1/S1Ex5F (5'-gctacggttctgttactcaca-3') and KIR3DL1/S1Ex5R (5'-tgaccaytcgtagggagagt-3'); KIR3DL1Ex7e9F (5'-agtggatcatcctcttcatc-3') and KIR3DL1Ex7e9R (5'-gtgtacaagatggatctgt-3'), resulting in fragments with 2582 bp, 1059 bp, 261 bp e 1912 bp, and 829 bp, respectively. Primers for amplification of exons 3 and 4 were adapted from Halfpenny and coworkers (2004) and for exons 1 and 2 obtained in collaboration with National Cancer Institute, Frederick, USA. The others were designed based on the human *KIR3DL1* (NG_005990.1) and *KIR3DS1* (NG_005994.1) genomic sequence. The amplicons were purified using polyethyleneglycol 8000 or exonuclease I and shrimp alkaline phosphatase enzymes and then, sequenced with DYEnamic ET Terminator Cycle Sequencing Kit in Megabace 1000 (GE Healthcare). Primers used for the sequencing reactions were: KIR3DL1/S1_EX1F (5'-tttagcatctcgtgttcgg-3'); and KIR3DL1/S1Ex2F (5'-agtgggggcagcaggggtg-3'); KIR3DL1/S1Ex3F (5'-ccatgccgtgtctactttgt-3') and KIR3DL1/S1Ex3R (5'-ctgtgaccatgatcacca-3'); for exon 4, the same primers used for PCR were used for sequencing; KIR3DL1/S1Ex5F (5'-gggagtgagttctcagctca-3') and

KIR3DL1/S1Ex5R (5'-tgaccactcgtagggagagtg-3'); KIR3DL1/S1Ex7F (5'-gtgcttgctckaaagagaygc-3') and KIR3DL1/S1Ex9F (5'-ctgcaggaacagaacagcc-3').

Computational and statistical analysis

Sequences were visualized and edited using Phred, Phrap and Consed softwares and aligned using Clustal-W multiple alignment, available in the Bioedit software. Identification of the gametic phase of two or more heterozygote sites was performed using the dbLRC (leukocyte receptor complex database) online software tool, available at NCBI (<http://www.ncbi.nlm.nih.gov/>). Allele and genotype frequencies for each variant site were determined by direct counting. The significance of the differences of allelic frequencies between populations was evaluated by the exact test of population differentiation using the ARLEQUIN 3.11 software package (Excoffier et al., 2005).

Cavalli-Sforza's chord measure of genetic distances from allelic and binding sites frequencies of *KIR3DL1/S1* was estimated with the PHYLIP package (Felsenstein, 2004) and used to construct the dendrograms by the Neighbor-Joining method (Saitou & Nei, 1987). Dendrograms were drawn using TreeView 1.6.6 (Page, 1996).

The direct estimate of per-site heterozygosity (H_{obs}) was derived from the average pairwise sequence difference (Nei, 1983) and compared to the expected per-site heterozygosity (H_{exp}), using the ARLEQUIN 3.11. This software package was also used to estimate Ewens-Watterson homozygosity (θ_W) (Ewens, 1972; Watterson, 1978). Tajima's D (D_T) (Tajima, 1989), Fu and Li's D and F (D_F and F) (Fu and Li, 1993) without an outgroup, and Fay and Wu's H (Fay & Wu, 2000) with the chimpanzee sequence (Genbank: **NC_006486**) as an outgroup were performed to test the null hypothesis of neutral evolution in the populations analyzed. The DnaSP v.5.10 software (Rozas et al., 2003) was used to calculate these parameters.

DnaSP was also used to estimate the number of synonymous and nonsynonymous sites and the number of synonymous and nonsynonymous substitutions. To test the hypothesis of selective neutrality, the codon-based test of neutrality averaged over all pairs of the 25 *KIR3DL1/S1* alleles included in MEGA version 4 was used (Tamura et al., 2007). These tests were done for the whole sequence and also for the region comprising exons 3, 4, and 5, as well as for the sequences coding for the predicted binding-site motifs in the exons 4 and 5, separately. For exons 1, 2, and 7, to compare the distribution of the non-

synonymous substitution per non-synonymous site (dN) and the synonymous substitution per synonymous site (dS), the RxC program employing the Metropolis algorithm was used for analysis of contingency tables to obtain unbiased estimates of exact P values (Miller, 1997).

Results

Extensive polymorphism of the *KIR3DL1/S1* gene was reported for all the populations analyzed in the present study. We found 25 alleles over all populations and most were common among them (Table 1), as well as among worldwide populations. We identified four new alleles differing from the previously described alleles at exons 3 and 5. These alleles were observed in particular populations. The three ancestral lineages of the *KIR3DL1/S1* gene (*005, *015, and 3DS1*013) were found in all populations and presented similar frequencies in Euro-Brazilians (30.4%, 33.3% and 24.5%, respectively). In Africans, lineage *015 is the most common and accounting for 78.2% of the alleles, while the 3DS1*013 lineage had low frequency in comparison to the other two lineages and to its frequency in non-African populations (4.3%). Conversely, we observed a reduced frequency of *005 lineage in Amerindians compared to the other two lineages and to the other populations. Its presence at low frequency in the Guarani M'byá and Guarani Kaiowá is most likely due to gene flow.

Africans were the most diverse, having the greatest number of alleles (16) and highest heterozygosity (Table 2). However, even Amerindians have also a high number of alleles and high heterozygosity that is most closely related to Asian-Brazilians (ASI-BR).

The allele frequencies differed significantly for all pairs of populations, even between the two Kaingang populations, which are the most closely related among the populations analyzed. In Amerindians and Asian-Br, allele frequencies were more broadly scattered than in the Euro-Br and Afro populations. The alleles *KIR3DS1*01301*, *KIR3DL1*005*, *01502 and *01501 occurred in populations of all continents. Alleles of the *004 group were common just to Africa and Europe human groups and *02901 to Asia and America, although some populations of these continents lack them. Alleles of the *017 and *002 groups seem to be restricted to Africa and Europe respectively. The occasional presence of these alleles in other continents is most likely due to gene flow.

The genetic distances revealed a correlation of allelic frequencies with geographic distribution of the populations, as showed by the dendrogram based on *KIR3DL1/S1* allelic frequencies. This dendrogram clusters Amerindians apart from populations of other ancestries and closest to Asian-Brazilians (Figure 2). When other populations were included (Figure 3), we can see four major continental groups, similar to the pattern observed in figure 2, with (a) Amerindians grouped together, being closely related to (b) populations from East Asia, (c) Europeans and Euro-descendents gathered together and (d) Africans.

Neutrality tests of the entire sequence of *KIR3DL1/S1* alleles in populations of this study showed no significant deviations from selective neutrality, although high values of Tajima's D test were observed in all populations, excluding Africans (Table 3). The same pattern was observed when analyzing separately exons that code for different domains of the *KIR3DL1/S1* receptor. No deviations from neutrality were observed comparing the proportions of non-synonymous substitutions per non-synonymous site (dN) with these of synonymous substitutions per synonymous site (dS) across the different regions of alleles identified in this study ($P > 0.05$ in all situations) (Table 4).

On the other hand, significant positive values of Tajima's D revealed that, in exon 4, the region comprising codons that recognize HLA molecules is under balancing selection in Amerindians ($P < 0.05$, data not show), diversifying residues for the interaction between these molecules. Nevertheless we cannot exclude the effect of demographic factors on the Tajima's D values.

Analyzing just codons 115-118, 138-141, 162-169 in exon 4 and codons 198-203, 225-230 e 276-283 in exon 5, that are predicted to directly contact with HLA (Norman et al., 2007), we observed an excess of non-synonymous substitutions compared to synonymous substitutions ($P = 0.006$, data not show), revealing positive selection. The topology of the dendrogram constructed based on the frequencies of motifs for HLA interaction (Figure 4) also indicates deviations from neutrality.

Discussion

The *KIR* genes are among the most polymorphic of the human genome and evolved rapidly after separation of the primates and other mammalian groups (Moretta and Moretta, 2004). The *KIR3DL1/S1* gene has 84 described alleles many of which are common in

populations and shared by the major human groups. It has been suggested that this gene evolved under balancing selection in humans for a long time (Norman et al., 2007).

All populations of this study showed significant differentiation. The distinctiveness of the Guarani and Kaingang populations has been previously reported (Petzl-Erler, Luz & Sotomaior, 1993; Tsuneto et al., 2003; Oliveira et al., 2008; Beltrame et al., 2011, submitted). In this study, differentiation was also observed between the two Kaingang populations, the most related among the analyzed populations. Africans presented the highest number of alleles (16) and the highest heterozygosity, despite being the smallest sample analyzed ($n=23$) as generally observed for most loci. Euro-Brazilians had also many alleles, as also seen for other European descendent populations (<http://www.allelefrequencies.net/>). Among the Amerindians, Guarani Nandeva had the highest genetic diversity, which agrees with the relatively high amount of gene flow (14%) from non-Amerindians into this population (Tsuneto et al., 2003). On the other hand, Guarani M'byá had the lowest diversity, in agreement with their genetic isolation and the effect of random genetic drift. The high genetic diversity for *KIR3DL1/S1* gene observed in Amerindians, although lower than that of the classical *HLA* loci, is similar to that of some other MHC genes, including *MICA*, *C4A*, *C4B* (Petzl-Erler & McDevitt, 1994; Weg-Remers et al., 1997; Tsuneto et al., 2003; Oliveira et al., 2008), contrasting with the comparatively low genetic diversity observed for other still polymorphic loci, such as *TAP*, *LMP*, *CD80* and *CD86* (Faucz, Probst & Petzl-Erler, 2000; Beltrame et al., 2011, submitted).

The distribution of *KIR3DL1/S1* alleles among the populations of this study is quite related to their history. Natives from America share the same recent ancestral populations. They originated from Eastern Asian migrations most probably during the last glaciation (about 20,000 years ago), occupying, first, North America, and later, Central and South America (Fagundes et al., 2008; Fagundes, Kanitz & Bonatto, 2008). Despite their recent origin and the fact that they were originated from a reduced effective population size, Amerindians genetic diversity is shaped mainly by genetic drift and demographic effects, such as founder effect and bottlenecks. The relative contribution of each of these factors depends on the genomic region analyzed.

The worldwide distribution of alleles *KIR3DS1*01301*, *KIR3DL1*005*, **01502* and **01501* is an evidence that they originated in the ancestral African population. In the present work, Amerindians had the lowest number of alleles and, in these populations, allele *KIR3DS1*01301* was the most common, as reported for other Amerindians (Norman et al., 2007). Some alleles have low frequencies in Amerindians and they were not shared among

all populations. These alleles are probably a result from gene flow with non Amerindians, such as allele *00401, found in a single Guarani individual.

The high frequency of the activating lineage (3DS1) in Amerindians is consistent with the conclusion of other studies that proposed that activating *KIR* genes were important for survival of these populations during human evolution, because of their role in innate immune system (Gendzekhadze et al., 2006; Flores et al., 2007; Rajalingam et al., 2008; Kulkarni et al. 2008). Furthermore, the absence of allelic group *004, that is not expressed on the NK cell surface, in Amerindians also is an evidence of the importance of activating function of *KIR* genes in these populations. In other populations studied, the lineage *004 is frequent (Pando et al., 2003), including the Africans and Euro-Brazilians from this study.

The occurrence of allele *02901 only in Eastern Asians and Amerindians is related to their evolutionary relationship, in which Amerindians were originated from Eastern Asian migrations. Furthermore, similarity of the number of alleles and their frequency between Amerindians and Asian-Brazilians are also an evidence of their evolutionary proximity.

In contrast to other populations, Africans had the lowest frequency for the 3DS1 lineage. The same was observed in other African populations and was interpreted as evidence of purifying selection against 3DS1 lineage in Africa (Norman et al., 2007). Euro-Brazilians presented similar frequencies of the three major *KIR3DL1/S1* lineages, as European populations (<http://www.allelefreqencies.net/>). This may be the result of balancing selection.

Recombination played an important role in *KIR* evolution and the rate of expansion of *KIR* gene family seems to be one of the highest among all hominoid gene families (Hao & Nei., 2005). Of the four new *KIR3DL1/S1* alleles identified in this study, three were originated from point mutations (*v.3DL1*309A*, *v.3DL1*740A* e *v.3DS1*857T*) revealing that this evolution factor is also important for *KIR* diversification. The other new allele (*v.3DS1*732C*), identified in Guarani, is, most probably, a result from a gene conversion event in the exon 5 between an allele of a 3DL1 lineage and other from 3DS1 lineage. One new allele was identified in Africans and three in Amerindians. The allele *v.3DL1*309A* is common among three Guarani populations and probably was originated before the separation of these groups or was shared by gene flow. The analysis of other Amerindians and Asian populations could help to understand the origin and dispersion of these new alleles in human populations.

The relationship among different human populations based on allele frequencies showed the same topology observed from neutral markers (Ramachandran et al., 2005), grouping together historically related populations. This is consistent with the results from neutrality tests, which did not reveal significant deviations from selective neutrality.

The non-significant values for neutrality tests were in contrast with evidences of balancing selection acting in most human populations around the world, except in Africans, in which there were evidences for directional selection, favoring alleles related to the *015 lineage and of purifying selection against the 3DS1 lineage, as reported by other recent study (Norman et al., 2007). However, also in the present study we observed high Tajima's D values in all populations except in Africans and the lack of statistical significance may be due to the small sample sizes.

When we focused on gene regions that encode for sites in exons 4 and 5, predicted to contact HLA molecules, we found significantly positive Tajima's D values for exon 4. Moreover, when we analyzed only the codons corresponding to the positions of the molecule predicted to contact the HLA molecules, we observed an excess of non-synonymous substitutions, showing that coevolution with HLA could lead to a diversification of these sites, as previously suggested (Hsu et al., 2002; Norman et al., 2007; Martinez-Borra e Khakoo, 2008). Also the topology of the dendrogram based on the frequencies of the binding motifs differed from that expected under neutral evolution, as seem also by Norman and coworkers (2007).

In summary, the *KIR3DL1/S1* gene is highly polymorphic in all populations studied, including Amerindians, and evidences of balancing natural selection were observed only in regions that bind the HLA molecules.

Acknowledgments

We thank the individuals who donated blood samples for collaboration in this study. The technical assistance and friendship of our laboratory colleagues are highly valued. This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação Araucária de Apoio ao Desenvolvimento Científico e Tecnológico do Paraná, and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Table 1. Allele frequencies of *KIR3DL1/S1* gene observed in this study.

| | Population | | | | | | | |
|---------------------|-------------------|-----------------------|----------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| | AFRI 2n=46 (%) | EURO-BR 2n=102 (%) | ASI-BR 2n=112 (%) | GKW 2n=128 (%) | GND 2n=116 (%) | GRC 2n=134 (%) | KIV 2n=120 (%) | KRC 2n=144 (%) |
| Lineage *3DS1 | | | | | | | | |
| *01301 | 2 (4.3) | 25 (24.5) | 24 (21.4) | 55 (43.0) | 38 (32.7) | 71 (53.0) | 54 (45.0) | 61 (42.4) |
| v.3DS1*732C | | | | 1 (0.8) | | | | |
| v.3DS1*857T | | | | | 2 (1.7) | | | |
| Lineage *005 | | | | | | | | |
| *00401 | 5 (10.9) | 14 (13.7) | | | 1 (0.9) | | | |
| *00402 | 1 (2.2) | 1 (1.0) | | | | | | |
| *00501 | 1 (2.2) | 14 (13.7) | 8 (7.1) | 2 (1.6) | 12 (10.3) | 1 (0.7) | 13 (10.8) | 21 (14.6) |
| *019 | | 1 (1.0) | | | | | | |
| *041 | 1 (2.2) | 1 (1.0) | | | | | | |
| Lineage *015 | | | | | | | | |
| *0020101 | | 12 (11.8) | 1 (0.9) | | | | | |
| *00701 | 3 (6.5) | 1 (1.0) | 12 (10.7) | | 1 (0.9) | | | |
| *0080101 | 1 (2.2) | 3 (2.9) | 1 (0.9) | 1 (0.8) | | | | |
| *01501 | 11 (23.9) | 8 (7.8) | 3 (2.7) | 1 (0.8) | 1 (0.9) | 19 (14.2) | 5 (4.2) | |
| *01502 | 5 (10.9) | 9 (8.8) | 48 (42.9) | 13 (10.1) | 14 (12.1) | 15 (11.2) | 40 (33.3) | 53 (36.8) |
| *016 | | | | | 1 (0.9) | | 1 (0.8) | |
| *01701 | 6 (13.0) | | | 1 (0.8) | | | | |
| *01702 | 2 (4.3) | | | | 1 (0.9) | | | |
| *018 | 2 (4.3) | | | | | | | |
| *020 | 3 (6.5) | 1 (1.0) | 5 (4.5) | | | | 4 (3.3) | 1 (0.7) |
| *025 | 1 (2.2) | | | | | | | |
| *02901 | | | 3 (2.7) | 46 (35.9) | 29 (25.0) | 26 (19.4) | 2 (1.7) | 7 (4.9) |
| *033 | 1 (2.2) | | | | | | | |
| v.3DL1*309A | | | | 1 (0.8) | 2 (1.7) | 2 (1.5) | | |
| v.3DL1*740A | 1 (2.2) | | | | | | | |
| Recombinant alleles | | | | | | | | |
| *00101 | | 11 (10.8) | 7 (6.3) | 7 (5.5) | 14 (12.1) | | 1 (0.8) | 1 (0.7) |
| *009 | | 1 (1.0) | | | | | | |

AFRI - Africans; EURO-BR - Euro-Brazilians; ASI-BR - Asian-Brazilians; GKW - Guarani Kaiowá; GND - Guarani Nandeva; GRC - Guarani M'byá; KIV - Kaingang from Ivaí; KRC - Kaingang from Rio das Cobras. In bold - alleles identified in this work.

Table 2. Number of alleles and heterozygosity values observed in each population.

| | AFRI (n=23) | EURO-BR (n=51) | ASI-BR (n=56) | GKW (n=64) | GND (n=58) | GRC (n=67) | KIV (n=60) | KRC (n=72) |
|------------------|----------------|-------------------|------------------|---------------|---------------|---------------|---------------|---------------|
| N | 16 | 14 | 10 | 10 | 12 | 6 | 8 | 6 |
| H _{obs} | 0.913 | 0.696 | 0.726 | 0.656 | 0.690 | 0.716 | 0.667 | 0.611 |
| H _{exp} | 0.904 | 0.871 | 0.746 | 0.678 | 0.796 | 0.654 | 0.684 | 0.666 |

N - Number of alleles; H_{obs} - Observed heterozygosity; H_{exp} - Expected heterozygosity; n - Number of individuals. AFRI - Africans; EURO-BR - Euro-Brazilians; ASI-BR - Asian-Brazilians; GKW - Guarani Kaiowá; GND - Guarani Nandeva; GRC - Guarani M'byá; KIV - Kaingang from Ivaí; KRC - Kaingang from Rio das Cobras.

Table 3. *KIR3DL1/S1* sequence diversity parameters of populations.

| Region | $\frac{d}{n}$ | AFRI | EURO-BR | ASI-BR | GKW | GND | GRC | KIV | KRC |
|---|----------------|--------|---------|--------|--------|--------|--------|--------|--------|
| Complete Coding Sequence | N | 46 | 102 | 112 | 128 | 116 | 134 | 120 | 144 |
| | Alleles | 16 | 14 | 10 | 10 | 12 | 6 | 8 | 6 |
| | S | 40 | 38 | 33 | 31 | 38 | 30 | 30 | 30 |
| | | 6.933 | 13.407 | 9.463 | 11.931 | 12.454 | 11.430 | 12.254 | 12.458 |
| | θ_w | 0.115 | 0.139 | 0.254 | 0.328 | 0.211 | 0.351 | 0.329 | 0.339 |
| | D_T | -0.680 | 2.588 | 1.569 | 3.232 | 2.283 | 3.194 | 3.576 | 3.806 |
| | D_F | 1.458 | 1.502 | 1.369 | 1.299 | 0.346 | 0.170 | 1.966 | 1.620 |
| | F | 0.833 | 2.320 | 1.737 | 2.489 | 1.356 | 1.418 | 3.110 | 3.006 |
| | H | -8.487 | 0.483 | -4.065 | -1.823 | -1.109 | -4.705 | -1.375 | -0.427 |
| 5' UTR and Exons 1 and 2 (Signal peptide) | S | 10 | 8 | 8 | 8 | 8 | 8 | 8 | 8 |
| | | 2.245 | 3.275 | 2.879 | 3.168 | 3.416 | 3.035 | 3.388 | 3.462 |
| | θ_w | 0.115 | 0.139 | 0.254 | 0.328 | 0.211 | 0.351 | 0.329 | 0.339 |
| | D_T | -0.038 | 2.727 | 2.157 | 2.689 | 3.025 | 2.500 | 2.999 | 3.220 |
| | D_F | 0.000 | 1.258 | 1.251 | 1.242 | 1.249 | -0.443 | 1.247 | 1.233 |
| | F | -0.337 | 2.085 | 1.854 | 2.071 | 2.208 | -0.689 | 2.198 | 2.291 |
| | H | -1.271 | 0.523 | -1.247 | 1.315 | -0.915 | -2.000 | -0.921 | -0.633 |
| | S | 20 | 17 | 20 | 16 | 16 | 16 | 20 | 18 |
| Exons 3, 4 and 5 (Ig-like domains) | | 3.644 | 5.335 | 6.019 | 4.891 | 5.570 | 5.805 | 6.774 | 4.382 |
| | θ_w | 0.115 | 0.328 | 0.211 | 0.351 | 0.329 | 0.339 | 0.139 | 0.254 |
| | D_T | -0.644 | 1.925 | 1.714 | 1.814 | 2.367 | 2.704 | 2.196 | 0.810 |
| | D_F | 0.956 | 0.606 | 0.362 | -1.128 | 1.620 | 1.061 | 0.844 | 0.703 |
| | F | 0.485 | 1.319 | 1.048 | -0.030 | 2.267 | 2.011 | 1.616 | 0.888 |
| | H | -2.709 | -0.684 | 1.133 | -2.383 | -0.357 | 0.294 | 1.569 | -0.442 |
| | S | 12 | 8 | 12 | 8 | 8 | 8 | 12 | 9 |
| Exon 7 (Transmembrane region) | | 0.157 | 3.969 | 3.732 | 4.016 | 3.993 | 3.934 | 4.111 | 2.911 |
| | θ_w | 0.115 | 0.328 | 0.211 | 0.351 | 0.329 | 0.339 | 0.139 | 0.254 |
| | D_T | -1.010 | 3.957 | 1.709 | 4.060 | 3.957 | 3.973 | 2.068 | 1.747 |
| | D_F | 1.477 | 1.242 | -0.468 | 1.239 | 1.247 | 1.234 | 1.467 | 1.312 |
| | F | 0.775 | 2.593 | 0.375 | 2.637 | 2.591 | 2.603 | 2.000 | 1.737 |
| | H | 1.399 | 0.630 | 1.812 | 0.390 | 0.524 | 0.735 | 0.579 | -0.592 |

N - number of chromosomes; S - number of segregating sites; d ($\times 10^{-4}$) - nucleotide diversity per site, θ_w ($\times 10^{-4}$) - Watterson's Theta per site from S; D_T - Tajima's D; D_F - Fu and Li's D without an outgroup (D^*); F - Fu and Li's F without an outgroup (F^*); H - Fay and Wu's H with the chimpanzee sequence (Genbank: **NC_006486**) as an outgroup. AFRI - Africans; EURO-BR - Euro-Brazilians; ASI-BR - Asian-Brazilians; GKW - Guarani Kaiowá; GND - Guarani Nandeva; GRC - Guarani M'byá; KIV - Kaingang from Ivai; KRC - Kaingang from Rio das Cobras.

Table 4. Comparison of synonymous substitutions per synonymous site and non-synonymous substitutions per non-synonymous site to test the hypothesis of selective neutrality.

| Region | Ns | Ms | Ms/Ns | Na | Ma | Ma/Na |
|--|-------|----|-------|-------|----|-------|
| Exons 1 and 2 (Signal peptide) | 14.8 | 0 | 0.000 | 54.2 | 3 | 0.055 |
| Exons 3, 4, and 5 (Ig-like domains) | 216.8 | 14 | 0.065 | 663.2 | 44 | 0.066 |
| Exon 7 (Transmembrane region) | 25.0 | 5 | 0.200 | 80.0 | 14 | 0.175 |

Ns: number of synonymous sites (average for all sequences); Ms: number of synonymous substitutions; Na: number of non-synonymous sites (average for all sequences); Ma: number of non-synonymous substitutions. $p > 0.05$ for all comparisons.



Figure 1. South America's map with the location of Curitiba and the reservation areas where the Guarani and Kaingang populations and from this study live (Oliveira et al., 2008). Note: 1 - Ivaí; 2 - Rio das Cobras; 3 - Amambaí and Limão Verde; 4: Porto Lindo; 5 - Curitiba.

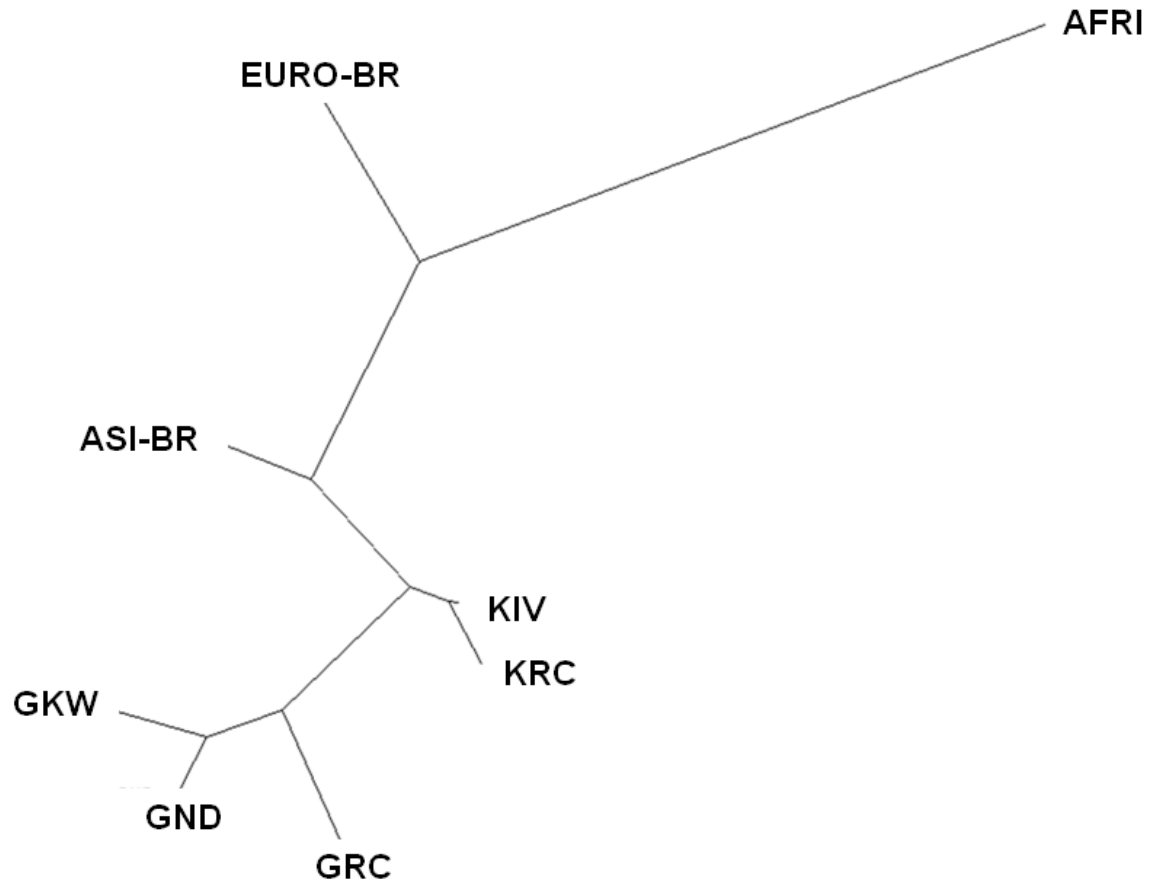


Figure 2. Dendrogram based on Cavalli-Sforza's chord measure of genetic distances from allelic frequencies for the *KIR3DL1/S1* gene in populations of this study. AFRI - Africans; EURO-BR - Euro-Brazilians; ASI-BR - Asian-Brazilians; GKW - Guarani Kaiowá; GND - Guarani Nandeva; GRC - Guarani M'byá; KIV - Kaingang from Ivaí; KRC - Kaingang from Rio das Cobras.

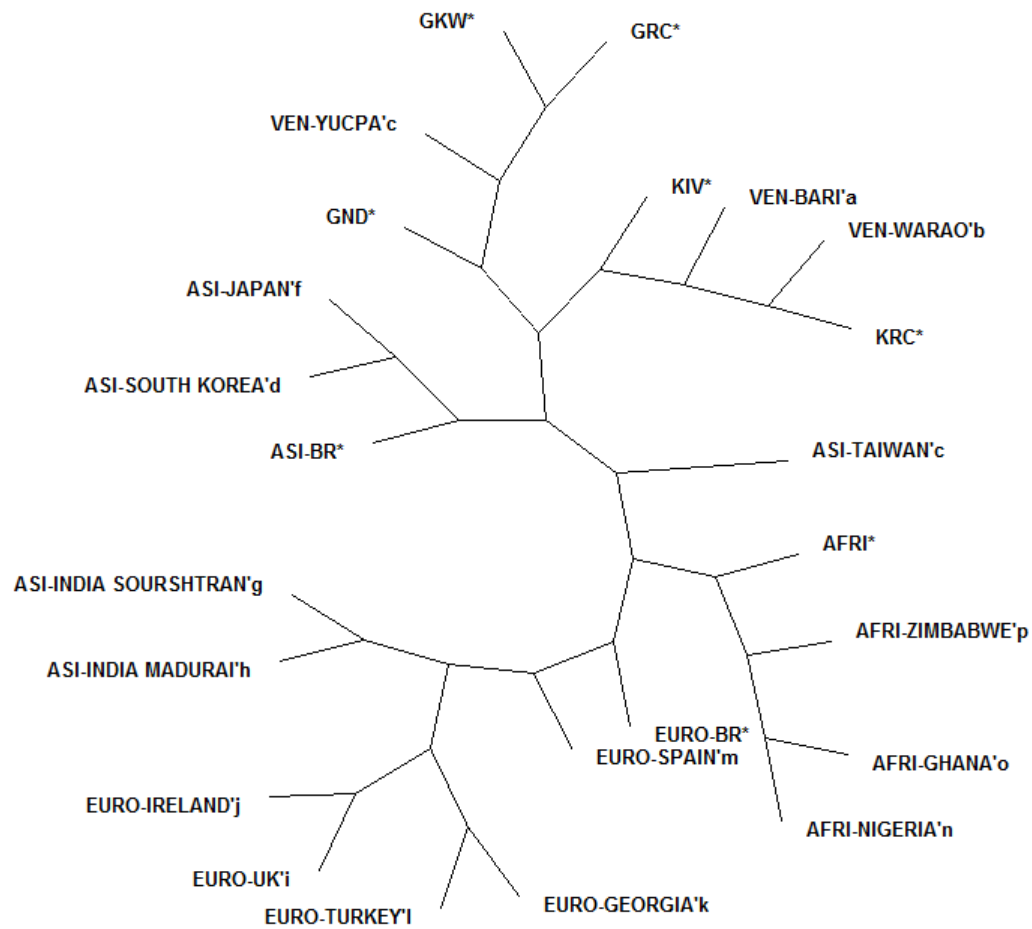


Figure 3. Dendrogram based on Cavalli-Sforza's chord measure of genetic distances from allelic frequencies for the *KIR3DL1/S1* gene in populations around the world. AFRI - Africans; EURO-BR - Euro-Brazilians; ASI-BR - Asian-Brazilians; GKW - Guarani Kaiowá; GND - Guarani Nandeva; GRC - Guarani M'byá; KIV - Kaingang from Ivaí; KRC - Kaingang from Rio das Cobras; a,b,c - Amerindians from Venezuela (Norman et al., 2007); d - Asians from South Korea (Whang et al., 2005); e - Asians from Taiwan (Norman et al., 2007); f - Asians from Japan (Yawata et al., 2006); g - Asians Sourshtran from India (Norman et al., 2007); h - Asians Madurai from India (Norman et al., 2007); i - Europeans from the United Kingdom (Norman et al., 2007); j - Europeans from Northern Ireland (Middleton, 2003); k - Europeans from Georgia (Norman et al., 2007); l - Europeans from Turkey (Norman et al., 2007); m - Europeans from Spain (Norman et al., 2007); n - Africans from Nigeria (Norman et al., 2007); o - Africans from Ghana (Norman et al., 2007); p - Africans from Zimbabwe (Norman et al., 2007).

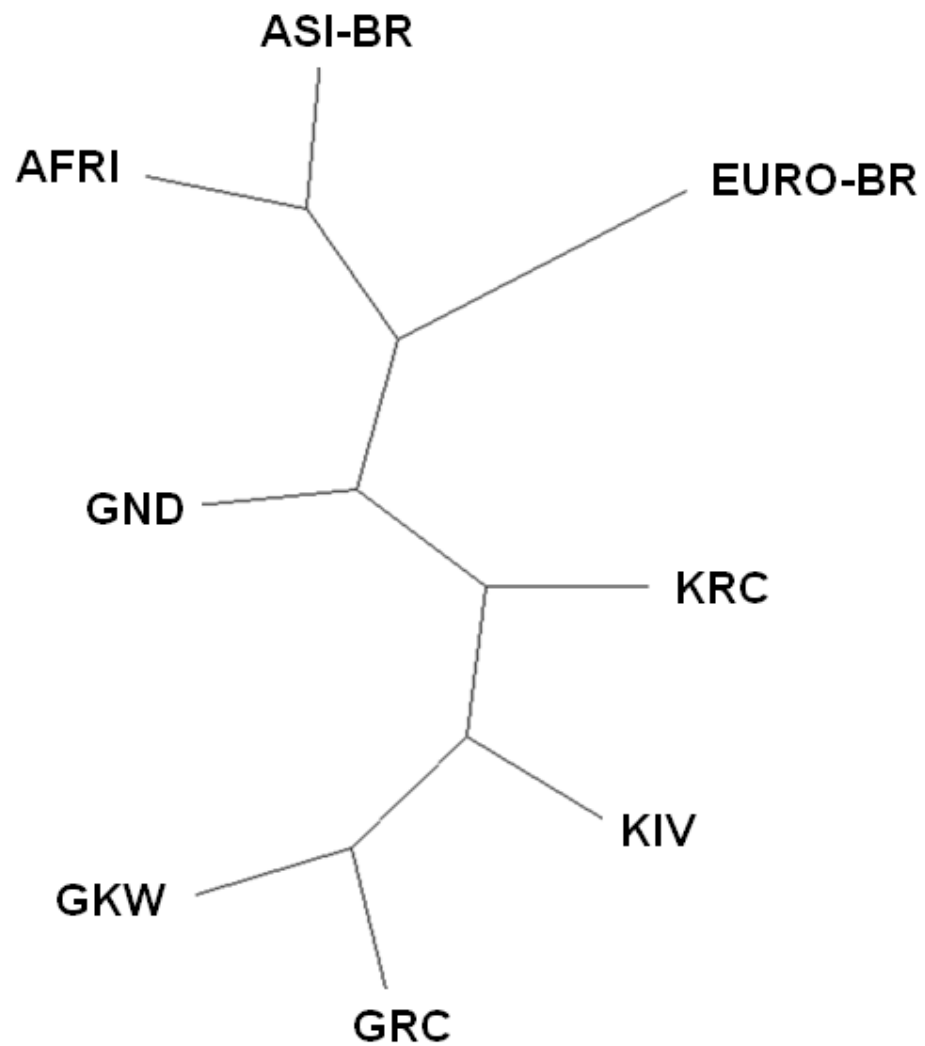


Figure 4. Dendrogram based on Cavalli-Sforza's chord measure of genetic distances from binding sites motif frequencies for the *KIR3DL1/S1* gene in populations of this study. AFRI - Africans; EURO-BR - Euro-Brazilians; ASI-BR - Asian-Brazilians; GKW - Guarani Kaiowá; GND - Guarani Nandeva; GRC - Guarani M'byá; KIV - Kaingang from Ivaí; KRC - Kaingang from Rio das Cobras.

References

- Ashworth, L. K.; Batzer, M. A.; Brandriff, B.; Branscomb, E.; De Jong, P.; Garcia, E.; Garnes, J. A.; Gordon, L. A.; Lamerdin, J. E.; Lennon, G.; Mohrenweiser, H.; Olsen, A. S.; Slezak, T. and Carrano, A. V. An integrated metric physical map of human chromosome 19. **Nature. Genetics**, v. 11, p. 422-427, 1995.
- Carr, W. H.; Pando, M. J. and Parham, P. KIR3DL1 Polymorphisms That Affect NK Cell Inhibition by HLA-Bw4 Ligand. **The Journal of Immunology**, v. 175, p. 5222-5229, 2005.
- Chan, H. W.; Kurago, Z. B.; Stewart, C. A.; Wilson, M. J.; Martin, M. P.; Mace, B. E.; Carrington, M. Trowsdale, J. and Lutz, C. T. DNA Methylation Maintains Allele-specific *KIR* Gene Expression in Human Natural Killer Cells. **The Journal of Experimental Medicine**, v. 197, p. 2245-2255, 2003.
- Denis, L.; Sivula, J.; Gourraud, P. A.; Kerdudou, N.; Chout, R.; Ricard, C.; Moisan, J. P.; Gagne, K.; Partanen, J. and Bignon, J. D. Genetic diversity of KIR natural killer cell markers in populations from France, Guadeloupe, Finland, Senegal and Reunion. **Tissue Antigens**, v. 66, p. 267-276, 2005.
- Di Santo, J. P. Natural Killer cell developmental pathways: A question of balance. **Annual Reviews Immunology**, v. 24, p. 257-286, 2006.
- Ewens, W. J. The sampling theory of selectively neutral alleles. **Theoretical Population Biology**, v. 3, p. 87-112, 1972.
- Excoffier, L.; Laval, G. and Schneider, S. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. **Evolutionary Bioinformatics Online**, v. 1, p. 47-50, 2005.
- Fagundes, N. J.; Kanitz, R.; Eckert, R.; Valls, A. C.; Bogo, M. R.; Salzano, F. M.; Smith, D. G.; Silva, W. A. Jr.; Zago, M. A.; Ribeiro-Dos-Santos, A. K.; Santos, S. E.; Petzl-Erler, M. L. and Bonatto, S. L. Mitochondrial population genomics supports a single pre-Clovis origin with a coastal route for the peopling of the Americas. **American Journal of Human Genetics**, v. 82, p. 583-592, 2008a.
- Fagundes, N. J. R.; Kanitz, R. and Bonatto, S. L. A reevaluation of the native American mtDNA genomic diversity and its bearing on the models of early colonization of Beringia. **PLOS One**, v. 3, 2008b.
- Faucz, F. R.; Probst, C. M. and Petzl-Erler, M. L. Polymorphism of LMP2, TAP1, LMP7 and TAP2 in Brazilian Amerindians and Caucasoids: implications for the evolution of allelic and haplotypic diversity. **European Journal of Immunogenetics**, v. 27, p. 5-16, 2000.
- Fay, J. C. and Wu, C. I. Hitchhiking under positive Darwinian selection. **Genetics**, v. 155, p. 1405-1413, 2000.
- Felsenstein, J. PHYLIP (Phylogeny Inference Package) version 3.6. Programa distribuido pelo autor. Department of Genome Sciences, University of Washington, Seattle, 2004.

Flores A. C.; Marcos C. Y.; Paladino N.; Capucchio, M.; Theiler, G.; Arruvito, L.; Pardo, R.; Habegger, A.; Williams, F.; Middleton, D. and Fainboim, L. KIR genes polymorphism in Argentinean Caucasoid and Amerindian populations. **Tissue Antigens**, v. 69, p. 568-576, 2007.

Fu, Y. X. and W. H. Li. Statistical tests of neutrality of mutations. **Genetics**, v. 133, p. 693-709, 1993.

Gardiner, C. M.; Killer cell immunoglobulin-like receptors on NK cells: the how, where and why. **International Journal of Immunogenetics**, v. 35, p. 1-8, 2007.

Gardiner, C. M.; Guethlein, L. A.; Shilling, H. G.; Pando, M.; Carr, W. H.; Rajalingam,; Vilches, R. C. and Parham, P. Different NK Cell Surface Phenotypes Defined by the DX9 Antibody Are Due to *KIR3DL1* Gene Polymorphism. **The Journal of Immunology**, v. 166, p. 2992-3001, 2001.

Gendzekhadze, K.; Norman, P. J.; Abi-Rached, L.; Layrisse, Z. and Parham, P. High KIR diversity in Amerindians is maintained using few gene-content haplotypes. **Immunogenetics**, v. 58, p. 474-480, 2006.

Halfpenny, I. A.; Middleton, D.; Barnett, Y. A. and Williams, F. Investigation of Killer Cell Immunoglobulin-Like Receptor Gene Diversity: IV. *KIR3DL1/S1*. **Human Immunology**, v. 65, p. 602-612, 2004.

Hao, L. and Nei, M. Rapid expansion of killer cell immunoglobulin-like receptor genes in primates and their coevolution with MHC class I genes. **Gene**, v. 347, p. 149-159, 2005.

Hiby, S. E.; Walker, J. J.; O'shaughnessy, K. M.; Redman, C. W.G.; Carrington, M.; Trowsdale, J.; Moffett, A. Combinations of maternal KIR and fetal HLA-C genes influence the risk of preeclampsia and reproductive success. **Journal of Experimental Medicine**, v. 200, p. 957-965, 2004.

Hsu, K. C.; Chida, S.; Geraghty, D. E. and Dupont, B. The killer cell immunoglobulin-like receptor (KIR) genomic region: gene order, haplotypes and allelic polymorphism. **Immunological Reviews**, v. 190, p. 40-52, 2002.

Hughes, A. L. Evolution of the human killer cell inhibitory receptor family. **Molecular Phylogenetics and Evolution**, v. 25p. 330-340, 2002.

Kim, S.; Sunwoo, J. B.; Yang, L.; Choi, T.; Song, Y. J.; French, A. R.; Vlahiotis, A.; Piccirillo, J. F.; Cella, M.; Colonna, M.; Mohanakumar, T.; Hsu, K. C.; Dupont, B. and Yokoyama, W. M. HLA alleles determine differences in human natural killer cell responsiveness and potency. **Proceedures of National Academy of Sciences**, v. 105, p. 3053-3058, 2008.

Kulkarni, S.; R. Single, M.; Martin, M. P.; Rajalingam R.; Badwe, R. Joshi, N. and Carrington, M. Comparison of the rapidly evolving KIR locus in Parsis and natives of India. **Immunogenetics**, v. 60, p. 121-129, 2008.

Li, H.; Pascal, V.; Martin, M. P.; Carrington, M. and Anderson, S. K. Genetic control of variegated *KIR* gene expression: polymorphisms of the bi-directional *KIR3DL1* promoter are associated with distinct frequencies of gene expression. **PLOS Genetics**, v. 4, 2008.

Martin, M. P.; Qi, Y.; Gao, X.; Yamada, E.; Martin, J. N.; Pereyra, F.; Colombo, S.; Brown, E. E.; Lesley, S. W.; Phair, J.; Goedert, J. J.; Buchbinder, S.; Kirk, D. G.; Telenti, A.; Connors, M.; O'Brien, S. J.; Walker, B. D.; Parham, P.; Deeks, S. G.; McVicar, D. W. and Carrington, M. Innate partnership of HLA-B and KIR3DL1 subtypes against HIV-1. **Nature Genetics**, v. 39, 2007.

Martinez-Borra, J. and Khakoo, S. I. Speed and selection in the evolution of killer-cell immunoglobulin-like receptors. **International Journal Immunogenetics**, v. 35, p. 89-96, 2008.

Middleton, D.; Meenagh, A. and Gourraud, P. A. KIR haplotype content at the allele level in 77 Northern Irish families. **Immunogenetics**, v. 59, p. 145-58, 2007.

Miller, M. P. RxC: a program for the analysis of contingency tables via the metropolis algorithm computer. **Department of Biological Sciences**, Northern Arizona University, Flagstaff, 1997.

Moretta, L. and Moretta, A. Killer immunoglobulin-like receptors. Current **Opinion in Immunology**, v. 16, p. 626-633, 2004.

Nei, M.; Tajima, F. and Tateno, Y. Accuracy of estimated phylogenetic trees from molecular data. **Journal of Molecular Evolution**, v. 19, p. 153-170, 1983.

Norman, P. J.; Abi-Rached, L.; Gendzekhadze, K.; Hammond, J. A.; Moesta, A. K.; Sharma, D.; Graef, T.; McQueen, K. L.; Guethlein, L. A.; Carrington, C. V. F.; Chandanayingyong, D.; Chang, Y. H.; Crespi, C.; Saruhan-Direskeneli, G.; Hameed, K.; Kamkamidze, G.; Koram, K. A.; Layrisse, Z.; Matamoros, N.; Mila, J.; Park, M. H.; Pitchappan, R. M.; Dan Ramdath, D.; Shiao, M. Y.; Stephens, H. A. F.; Struik, S.; Tyan, D.; Verity, D. H.; Vaughan, R. W.; Davis, R. W.; Fraser, P. A.; Riley, E. M.; Ronaghi, M. and Parham, P. Meiotic Recombination generates rich diversity in NK cell receptor genes, alleles, and haplotypes. **Genome Research**, v. 19, p. 757-769, 2009.

Norman, P. J.; Abi-Rached, L.; Gendzekhadze, K.; Korb, D.; Gleimer, M.; Rowley, D.; Bruno, D.; Carrington, C. V. F.; Chandanayingyong, D.; Chang, Y-H.; Crespi, C.; Saruhan-Direskeneli, G.; Fraser, P. A.; Hameed, K.; Kamkamidze, G.; Koram, K. A.; Layrisse, Z.; Matamoros, N.; Mila, J.; Park, M. H.; Pitchappan, R. M.; Dan Ramdath, D.; Shiao, M-Y.; Stephens, H. A. F.; M.; Struik, S.; Verity, D. H.; Vaughan, R. W.; Tyan, D.; Davis, R. W.; Riley, E. M.; Ronaghi, M. and Parham, P. Unusual selection on the KIR3DL1/S1 natural killer cell receptor in Africans. **Nature Genetics**, v. 39, 2007.

O'Connor, G. M.; Guinan, K. J.; Cunningham, R. T.; Middleton, D.; Parham, P. and Gardiner, C. M. Functional Polymorphism of the KIR3DL1/S1 Receptor on Human NK Cells. **The Journal of Immunology**, v. 178, p. 235-241, 2007.

Oliveira, L.A.; Ribas, F.; Bicalho, M.G.; Tsuneto, L.T. and Petzl-Erler, M.L. High frequencies of alleles MICA*020 and MICA*027 in Amerindians and evidence of positive selection on exon 3. **Genes and Immunity**, v. 9, p. 697-705, 2008.

Page, R. D. M. Treeview: An application to display phylogenetic trees on personal computers. **Computer Applications in the Biosciences**, v. 12, p.357-358, 1996.

Pando, M. J.; Gardiner, C. M.; Gleimer, M.; McQueen, K. L. and Parham, P. The Protein Made from a Common Allele of *KIR3DL1*(3DL1*004) Is Poorly Expressed at Cell Surfaces due to Substitution at

Positions 86 in Ig Domain 0 and 182 in Ig Domain. **The Journal of Immunology**, v. 171, p. 6640-6649, 2003.

Parham, P. Taking license with natural killer cell maturation and repertoire development. **Immunological Reviews**, v.214, p. 155-160, 2006.

Petzl-Erler, M.L.; Luz, R. and Sotomaior, V.S. The HLA polymorphism of two distinctive South-American Indian tribes: The Kaingang and the Guarani. **Tissue Antigens**, v. 41, p. 227-237, 1993.

Petzl-Erler, M.L. and Mcdevitt, H.O. Molecular analysis of the HLA-DRB genes in two tribes of Brazilian Indians. **Human Immunology**, v. 41, p.180-184, 1994.

Rajalingam, R.; Du Z.; Meenagh, A.; Luo, L.; Kavitha, V. J.; Pavithra-Arulvani, R.; Vidhyalakshmi, A.; Sharma, S. K.; Balazs, I.; Reed, E. F.; Pitchappan, R. M. and Middleton, D. Distinct diversity of KIR genes in three southern Indian populations: comparison with world populations revealed a link between KIR gene content and pre-historic human migrations. **Immunogenetics**, v. 60, p. 207-217, 2008.

Rajalingam, R.; Parham, P. and Abi-Rached, L. Domain shuffling has been the main mechanism forming new hominoid killer cell Ig-like receptors. **The Journal of Immunology**, v. 172, p. 356-359, 2004.

Ramachandran, S.; Deshpande, O.; Roseman, C. C.; Rosenberg, N. A.; Feldman, M. W. and Cavalli-Sforza, L. L. Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa. **Proceedures of National Academy of Science of the United States of America**, v. 102, p. 15942-15947, 2005.

Rozas, J.; Sánchez Delbarrio, J. C.; Messeguer, X. and Rozas, R. DnaSP, DNA polymorphism analyses by the coalescent and other methods. **Bioinformatics**, v.19 p. 2496-2497, 2003.

Saitou, N. and Nei, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. **Molecular Biology and Evolution**, v. 4, p. 406–425, 1987.

Shilling, H. G.; Guethlein, L. A.; Cheng, N. W.; Gardiner, C. M.; Rodriguez, R.; Tyan, D. and Parham, P. Allelic polymorphism synergizes with variable gene content to individualize human KIR genotype. **The Journal of Immunology**, v. 168, p. 2307-2315, 2002.

Single, R. M.; Martin, M. P.; Gao, X.; Meyer, D.; Yeager, M.; Kidd, J. R.; Kidd, K. K. and Carrington, M. Global diversity and evidence for coevolution of KIR and HLA. **Nature Genetics**, 2007.

Tajima, F. Statistical methods to test for nucleotide mutation hypothesis by DNA polymorphism. **Genetics**, v. 123, p. 585-595, 1989.

Tamura, K.; Dudley, J.; Nei, M. and Kumar, S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. **Molecular Biology and Evolution**, v.24, p.1596-1599, 2007.

Thananchai, H.; Gillespie, G.; Martin, M. P.; Bashirova, A.; Yawata, N.; Yawata, M.; Easterbrook, P.; Mcvicar, D. W.; Maenaka, K.; Parham, P.; Carrington, M.; Dong, T. and Rowland-Jones, S. Cutting Edge: Allele-Specific and Peptide-Dependent Interactions between KIR3DL1 and HLA-A and HLA-B. **The Journal of Immunology**, v. 178, p. 33-37, 2007.

Thomas, R.; Yamada, E.; Alter, G.; Martin, M. P.; Bashirova, A. A.; Norman, P. J.; Altfeld, M.; Parham, P.; Anderson, S. K.; Mcvicar, D. W. and Carrington, M. Novel *KIR3DL1* Alleles and Their Expression Levels on NK Cells: Convergent Evolution of KIR3DL1 Phenotype Variation? **The Journal of Immunology**, v. 180, p. 6743-6750, 2008.

Toneva, M.; Lileapfaayge, V. G.; Dulphy, N.; Busson, M.; Lester, S.; Vu-Trien, A.; Michaylova, A.; Naumova, E.; McCluskey, J. and Charron, D. Brief communication: Genomic diversity of natural killer cell receptor genes in three populations. **Tissue Antigens**, v. 57, p. 358-362, 2001.

Trinchieri, G. Natural killer cells wear different hats: effector cells of innate resistance and regulatory cells of adaptive immunity and hematopoiesis. **Seminars in Immunology**, v. 7, p. 83-88, 1995.

Trowsdale, J. Genetic and functional relationship between MHC and NK receptor genes. **Immunity**, v. 15, p. 363-374, 2001.

Tsuneto, L. T.; Probst, C. M.; Hutz, M. H.; Salzano, F. M.; Rodriguez-Delfin, L. A.; Zago, M. A.; Hill, K.; Hurtado, A. M.; Ribeiro-Dos Santos, A. K. C. and Petzl-Erler, M. L. HLA class II diversity in seven Amerindian populations. Clues about the origins of the Aché. **Tissue Antigens**, v. 62, p. 512-526, 2003.

Watterson, G. The homozygosity test of neutrality. **Genetics**, v. 88, p. 405-417, 1978.

Weg-Remers, S.; Brenden, M.; Schwarz, E.; Witzel, K.; Schneider, P. M.; Guerra, L. K.; Rehfeldt, I. R.; Lima, M. T.; Hartmann, D.; Petzl-Erler, M. L.; De Messias, I. J. and Mauff, G. Major histocompatibility complex (MHC) class III genetics in two Amerindian tribes from Southern Brazil: the Kaingang and the Guarani. **Human Genetics**, v. 100, p. 548-556, 1997.

Wilson, M. J.; Torkar, M.; Haude, A.; Milne, S.; Jones, T.; Sheer, D.; Beck, S. and Trowsdale, J. Plasticity in the organization and sequences of human KIRyILT gene families. **Proceedings of National Academy of Sciences of the United States of America**, v. 97, p. 4778-4783, 2000.

Yawata, M. N.; Yawata, N.; Draghi, M.; Little, A.-M.; Partheniou, F. and Parham, P. Roles for HLA and KIR polymorphisms in natural killer cell repertoire selection and modulation of effector function. **The Journal of Experimental Medicine**, v. 203, p. 633-645, 2006.

cell repertoires of KIR expression. **Immunogenetics**, v. 54, p. 543-550, 2002.

Yu, J.; Heller, G.; Chewning, J.; Kim, S.; Yokoyama, W. M. and Hsu, K. C. Hierarchy of the Human Natural Killer Cell Response Is Determined by Class and Quantity of Inhibitory Receptors for Self-HLA-B and HLA-C Ligands. **The Journal of Immunology**, v. 179, p. 5977-5989, 2007.

5 RESULTADOS

5.1 METODOLOGIA DE TIPAGEM E MÉTODOS DE INFERÊNCIA HAPLOTÍPICA

Foram realizados vários testes para a identificação dos genótipos presentes em cada indivíduo através do programa *dbLRC*, com a finalidade de testar a sua confiabilidade. Nesses testes foram utilizados genótipos formados por alelos já descritos que levam a ambiguidades e também genótipos/alelos cujas sequências permitem inferir apenas um único genótipo.

Não houve ambiguidades na identificação dos alelos *3DL1*, com exceção de um indivíduo na população africana, para o qual a fase dos polimorfismos presentes nos exons 1 e 5 precisa ser conhecida para que o genótipo seja determinado. Por outro lado houve ambiguidades para a determinação do genótipo dos indivíduos portadores de pelo menos um alelo *3DS1*. As sequências obtidas são comuns aos alelos *KIR3DS1*01301*, **01302* e **045*. Estes alelos diferem na extremidade 3' do exon 5, uma região que não foi sequenciada neste estudo devido a localização dos oligonucleotídeos iniciadores para a amplificação deste exon. Os alelos *KIR3DS1*01302* e **045* foram descritos apenas em uma população cada, de origem européia e do leste asiático, respectivamente, nas quais suas frequências são baixas. Já o alelo *KIR3DS1*01301* foi identificado em inúmeras populações de diferentes etnias e apresenta frequência bastante elevada na maioria das populações estudadas, chegando a 51,6% nos ameríndios Yucpa da Venezuela. Desta forma, levando-se em consideração as frequências já descritas para os alelos *3DS1*, o alelo **01301* foi considerado como mais provável, nas populações deste estudo.

Para testar a confiabilidade da identificação genotípica através do *dbLRC*, foi feita a inferência haplotípica na população Guaraní M'byá (GRC), com o auxílio do programa ARLEQUIN (v 3.11). Para esta análise, foram considerados apenas os locos (sítios) informativos ao longo das sequências encontradas nesta população, com fase gamética desconhecida. Foram utilizados os algoritmos ELB (análise Bayesiana) e EM (análise por máxima verossimilhança), obtendo-se, respectivamente, 12 e 6 haplótipos (Tabela 2). Os seis haplótipos inferidos pelo algoritmo EM correspondem aos seis alelos identificados nesta população, através do programa *dbLRC*. O algoritmo ELB inferiu seis haplótipos adicionais que provavelmente são espúrios, apesar de alguns apresentarem frequência relativamente alta (por ex. haplótipo 7, com frequência de 6,7%).

TABELA 2. Comparação entre as frequências haplotípicas (alélicas) entre os diferentes métodos de inferência (2n=134 cromossomos).

| Haplótipo/Alelo | Freq. <i>dbLRC</i> | Freq. <i>EM</i> | Freq. <i>ELB</i> |
|--------------------|--------------------|-----------------|------------------|
| Hap1 (3DS1*01301) | 71 (53,0) | 71 (53,0) | 61 (45,5) |
| Hap2 (3DL1*02901) | 26 (19,4) | 26 (19,4) | 19 (14,2) |
| Hap3 (3DL1*01501) | 19 (14,2) | 18 (13,4) | 16 (11,9) |
| Hap4 (3DL1*01502) | 15 (11,2) | 16 (11,9) | 15 (11,2) |
| Hap5 (v.3DL1*309A) | 2 (1,5) | 2 (1,5) | 2 (1,5) |
| Hap6 (3DL1*00501) | 1 (0,7) | 1 (0,7) | 1 (0,7) |
| Hap7 | | | 9 (6,7) |
| Hap8 | | | 6 (4,5) |
| Hap9 | | | 2 (1,5) |
| Hap10 | | | 1 (0,7) |
| Hap11 | | | 1 (0,7) |
| Hap12 | | | 1 (0,7) |

dbLRC - *Leukocyte Receptor Complex database*; *EM* - análise por máxima verossimilhança; *ELB* - análise pseudo-Bayesiana. () - Frequência relativa.

A concordância entre os resultados de interpretação através do *dbLRC*, e da inferência haplotípica através do algoritmo *EM*, verificada em uma das amostras populacionais deste estudo, dá mais segurança para os resultados obtidos pelo programa *dbLRC*. Observa-se, contudo, que alguns métodos de inferência podem levar a interpretações errôneas, devendo os resultados ser interpretados com cautela. Assim, empregando-se o algoritmo *ELB* foram inferidos 12 haplótipos na população Guarani M'byá, sendo que apenas seis foram confirmados. Essas diferenças devem-se aos diferentes métodos empregados por cada algoritmo. A análise pseudo-Bayesiana (algoritmo *ELB*) é mais indicada para a determinação da fase gamética de sequências recombinantes, já o algoritmo *EM* (maximização de expectativa, por análise de máxima verossimilhança) é ideal para situações com um grande número de sítios polimórficos, desconsiderando a recombinação.

Os resultados obtidos neste estudo foram confrontados com resultados de genotipagem de presença e ausência para o gene *KIR3DL1/S1* obtidos em estudos anteriores do nosso grupo (AUGUSTO et al., manuscrito em preparação). Entre todos os indivíduos analisados, 19 (4,2%) apresentaram resultados discordantes entre os estudos.

A geração de dados pela técnica de sequenciamento, além de permitir a identificação de novas mutações, é bastante confiável, pois dispomos de programas que permitem verificar a qualidade de todas as bases que constituem a sequência em estudo. Os resultados servem como parâmetro para a elaboração e a avaliação de outros métodos

moleculares, empregados no estudo dos mesmos genes, como por exemplo, a PCR-SSP (*Polymarase Chain Reaction - Sequence Specific Primer*) e PCR-SSOP (*Polymerase Chain Reaction - Sequence Specific Oligonucleotide Probe*), que são muito comuns para a genotipagem de presença e ausência dos genes *KIR*. A maior parte dos indivíduos analisados neste estudo já havia sido genotipada para presença/ausência do gene *KIR3DL1/S1* e a discordância de 4,2% entre os resultados de sequenciamento e dos demais métodos sugere que os resultados obtidos pelas técnicas de PCR-SSP e SSOP devem ser interpretados com precaução, especialmente quando perfis de genes *KIR* com frequência bastante reduzida são inferidos pela presença ou ausência do gene *KIR3DL1/S1*.

5.2 DIVERSIDADE ALÉLICA E GENOTÍPICA

KIR3DL1/S1 é um dos genes *KIR* mais bem estudados. São conhecidos 84 alelos, que diferem por polimorfismos presentes ao longo de toda sequência do gene. Neste estudo, o genótipo *KIR3DL1/S1* presente em cada indivíduo foi identificado com o auxílio do programa *online dbLRC* (NCBI) a partir das sequências obtidas para os exons 1-5, 7 e 9.

Foram identificados 25 alelos considerando todas as populações analisadas neste estudo (Tabela 3). As frequências desses alelos diferem significativamente entre todas as populações, como evidenciado pelo teste exato de diferenciação populacional ($p < 10^{-5}$ para todas as comparações).

TABELA 3. Frequências alélicas para o gene *KIR3DL1/S1* observadas em casa população deste estudo.

| | População | | | | | | | |
|----------------------|-----------------|---------------------|--------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| | AFRI (2n=46) | EURO-BR (2n=102) | ASI-BR (2n=112) | GKW (2n=128) | GND (2n=116) | GRC (2n=134) | KIV (2n=120) | KRC (2n=144) |
| Linhagem *3DS1 | | | | | | | | |
| *01301 | 2 (4,3) | 25 (24,5) | 24 (21,4) | 55 (43,0) | 38 (32,7) | 71 (53,0) | 54 (45,0) | 61 (42,4) |
| v. 3DS1*732C | | | | 1 (0,8) | | | | |
| v. 3DS1*857T | | | | | 2 (1,7) | | | |
| Linhagem *005 | | | | | | | | |
| *00401 | 5 (10,9) | 14 (13,7) | | | 1 (0,9) | | | |
| *00402 | 1 (2,2) | 1 (1,0) | | | | | | |
| *00501 | 1 (2,2) | 14 (13,7) | 8 (7,1) | 2 (1,6) | 12 (10,3) | 1 (0,7) | 13 (10,8) | 21 (14,6) |
| *019 | | 1 (1,0) | | | | | | |
| *041 | 1 (2,2) | 1 (1,0) | | | | | | |
| Linhagem *015 | | | | | | | | |
| *00201 | | 12 (11,8) | 1 (0,9) | | | | | |
| *00701 | 3 (6,5) | 1 (1,0) | 12 (10,7) | | 1 (0,9) | | | |
| *00801 | 1 (2,2) | 3 (2,9) | 1 (0,9) | 1 (0,8) | | | | |
| *01501 | 11 (23,9) | 8 (7,8) | 3 (2,7) | 1 (0,8) | 1 (0,9) | 19 (14,2) | 5 (4,2) | |
| *01502 | 5 (10,9) | 9 (8,8) | 48 (42,9) | 13 (10,1) | 14 (12,1) | 15 (11,2) | 40 (33,3) | 53 (36,8) |
| *016 | | | | | 1 (0,9) | | 1 (0,8) | |
| *01701 | 6 (13,0) | | | 1 (0,8) | | | | |
| *01702 | 2 (4,3) | | | | 1 (0,9) | | | |
| *018 | 2 (4,3) | | | | | | | |
| *020 | 3 (6,5) | 1 (1,0) | 5 (4,5) | | | | 4 (3,3) | 1 (0,7) |
| *025 | 1 (2,2) | | | | | | | |
| *02901 | | | 3 (2,7) | 46 (35,9) | 29 (25,0) | 26 (19,4) | 2 (1,7) | 7 (4,9) |
| *033 | 1 (2,2) | | | | | | | |
| v. 3DL1*309A | | | | 1 (0,8) | 2 (1,7) | 2 (1,5) | | |
| v. 3DL1*740A | 1 (2,2) | | | | | | | |
| Alelos recombinantes | | | | | | | | |
| *00101 | | 11 (10,8) | 7 (6,3) | 7 (5,5) | 14 (12,1) | | 1 (0,8) | 1 (0,7) |
| *009 | | 1 (1,0) | | | | | | |

AFRI - Africanos; EURO-BR - Eurobrasileiros; ASI-BR - Brasileiros descendentes de Asiáticos; GKW - Guarani Kaiowá; GND - Guarani Nandeva; GRC - Guarani M'byá; KIV - Kaingang de Ivaí; KRC - Kaingang de Rio das Cobras. Entre parênteses são mostradas as frequências relativas para cada alelo, como porcentagens. Destacados em cinza os novos alelos identificados neste estudo.

Alguns alelos são exclusivos de determinadas populações, como, por exemplo, os alelos *3DL1*018*, **025* e **033* na população africana (Tabela 4). Entretanto a maioria dos alelos é compartilhada entre duas ou mais populações. Os alelos *3DS1*01301* e *3DL1*01502* possuem frequências elevadas (maior que 10%) em todas as populações, com exceção da população africana. A partir das frequências alélicas observadas no conjunto das populações analisadas neste estudo, foi construída uma rede mutacional dos alelos, com o auxílio do programa NETWORK (v. 4.6.0.0) (Figura 13). Nesta rede pode-se observar a relação entre os alelos, sendo que alelos mais próximos apresentam maior similaridade na sua sequência de nucleotídeos, e a distribuição das frequências alélicas nas populações estudadas.

A população africana (AFRI) apresentou o maior número de alelos: 16, seguida pela população euro-brasileira (EURO-BR), com 14 alelos. O grupo de alelos *004, que codifica proteínas que não são expressas na superfície das células, foi identificado apenas nas populações africana e EURO-BR e, em heterozigose com um alelo que é expresso normalmente, em um indivíduo Guarani Nandeva. Em todas as demais populações a linhagem *004 não foi detectada. Na população africana, dois alelos da linhagem *004 ocorreram em um mesmo indivíduo.

De forma geral, as populações ameríndias apresentaram um menor número de alelos e uma heterozigosidade reduzida, em comparação com as demais populações (Tabela 4).

TABELA 4. Número de alelos e heterozigosidade nas diferentes populações.

| | AFRI (n=23) | EURO-BR (n=51) | ASI-BR (n=56) | GKW (n=64) | GND (n=58) | GRC (n=67) | KIV (n=60) | KRC (n=72) |
|------------------|----------------|-------------------|------------------|---------------|---------------|---------------|---------------|---------------|
| N | 16 | 14 | 10 | 10 | 12 | 6 | 8 | 6 |
| H _{obs} | 0,913 | 0,696 | 0,726 | 0,656 | 0,690 | 0,716 | 0,667 | 0,611 |
| H _{esp} | 0,904 | 0,871 | 0,746 | 0,678 | 0,796 | 0,654 | 0,684 | 0,666 |

N - Número de alelos; H_{obs} - Heterozigosidade observada; H_{esp} - Heterozigosidade esperada. AFRI - Africanos; EURO-BR - Eurobrasileiros; ASI-BR - brasileiros descendentes de asiáticos; GKW - Guarani Kaiowá; GND - Guarani Nandeva; GRC - Guarani M'byá; KIV - Kaingang de Ivaí; KRC - Kaingang de Rio das Cobras.

As três linhagens ancestrais do gene *KIR3DL1/S1* (*005, *015 e 3DS1*013), que agrupam alelos com sequências similares e que codificam os mesmos motivos de ligação à HLA (ver figura 20), foram identificadas em todas as populações, embora suas frequências variem entre os grupos. A linhagem 3DS1*013 possui menor frequência na população africana e as frequências mais elevadas foram observadas nas populações ameríndias. Na população africana, a linhagem *015 representa 78,2% dos alelos. Nas populações ameríndias, a maior frequência da linhagem *015 é observada nos GKW (49,2%). Nos EURO-BR, as três linhagens possuem frequências mais bem distribuídas (30,4% dos alelos estão relacionadas a linhagem *005, 33,3% à linhagem *015 e 24,5% à linhagem 3DS1*013). Na população GRC apenas um alelo da linhagem *005 foi identificado. Essa linhagem também apresenta frequência reduzida na população GKW (1,6%), porém, ocorre em todas as demais populações com frequências entre 7,1 e 30,4%, sendo a frequência mais elevada observada em EURO-BR, na qual atinge 30,4%. A linhagem recombinante *001 é comum à maioria das populações e atinge frequência relativamente alta nos

indígenas GND (12,1%) e nos EURO-BR (10,8%). Outros alelos possuem frequência reduzida e ocorrem em grupos isolados.

A análise da variância molecular (AMOVA) mostra que a maior parte da variação genética deste loco ocorre dentro das populações (91,7%) e apenas 6,7% entre os diferentes grupos (Tabela 5). Essa variação é ainda menor quando observada entre populações de um mesmo grupo (1,8%).

TABELA 5. Análise da variância molecular (AMOVA) para o gene *KIR3DL1/S1* nas populações deste estudo.

| Fonte de Variação | Graus de Liberdade | Porcentagem da Variação |
|---|--------------------|-------------------------|
| Entre Grupos | 4 | 6,7 |
| Entre Populações - Dentro dos Grupos | 3 | 1,6 |
| Dentro das Populações | 894 | 91,7 |

Grupos: [AFRI], [EURO-BR], [ASI-BR], [GKW, GND e GRC] e [KIV e KRC].

Quatro novos alelos foram identificados nas populações deste estudo (Figura 14). No códon 82, exon 3, ocorre uma transversão (309C>A) (*v.3DL1*309A*), identificada apenas nas populações Guarani (GKW, GND e GRC), que não ocasiona substituição de aminoácidos na proteína. No exon 5 foram observadas três substituições de nucleotídeos ainda não descritas na literatura. Uma transição T>C (*v.3DS1*732C*) foi identificada no códon 223, caracterizando um novo alelo pertencente ao grupo de alelos ativadores (*3DS1*) em um indivíduo da população Guarani Kaiowá. Essa transição não ocasiona substituição de aminoácidos na proteína. A presença do nucleotídeo citosina (C) nessa posição é comum a vários alelos do grupo *3DL1*, o que sugere que este alelo tenha sido originado por um evento de conversão gênica. Uma transição de G>A foi identificada no códon 226, na população africana (*v.3DL1*740A*). Essa transição leva à substituição do aminoácido Arginina (Arg) por Glutamina (Gln) na sequência de aminoácidos e ocorre dentro de um dos sítios de interação do segundo domínio extracelular (D2) com as moléculas HLA. Ainda no exon 5, no códon 265, foi constatada uma outra transição (857C>T) (*v.3DS1*857T*), que ocasiona a substituição do aminoácido Treonina (Thr) por Isoleucina (Ile). Esse novo alelo foi identificado em dois indivíduos da população Guarani Nandeva. Os novos alelos foram confirmados através de uma nova reação de sequenciamento, utilizando um produto de PCR originário de outra reação de amplificação.

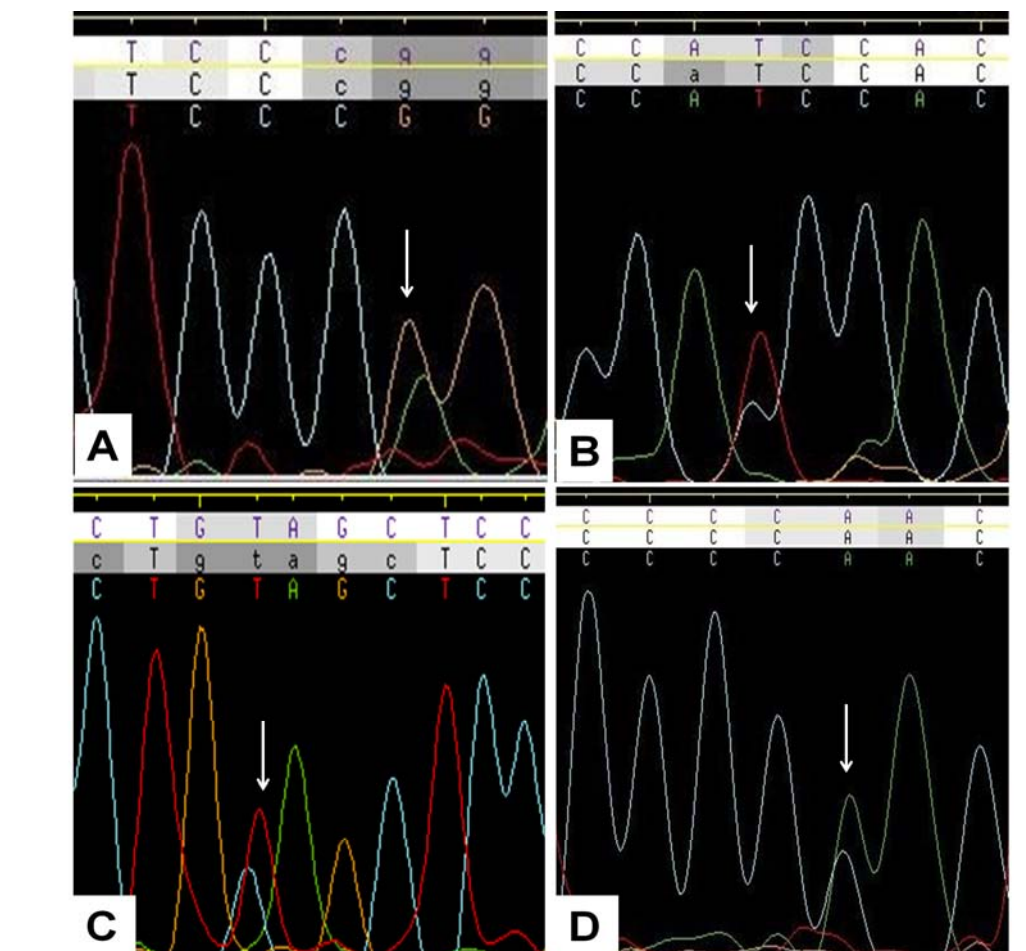


FIGURA 14. Cromatogramas mostrando as sequências identificadas para os novos alelos. A - Alelo do exon 3 (*v.3DL1*309A*) identificado nas populações Guarani; B - Alelo do exon 5 encontrado na população Guarani Kaiowá (*v.3DS1*732C*); C - Alelo identificado na população africana (*v.3DL1*740A*) e; D - Alelo *3DS1* identificado na população Guarani Nandeva (*v.3DS1*857T*). As setas indicam os picos dos heterozigotos onde foram identificados os novos alelos.

5.3 DIVERSIDADE ESTRUTURAL E FUNCIONAL DO PRODUTO GÊNICO

Os 25 alelos identificados neste estudo codificam 20 diferentes proteínas (Figura 15). Porém, quando são considerados apenas os resíduos de aminoácidos localizados no sulco de interação com os epítomos das moléculas HLA, essas proteínas formam cinco diferentes motivos (Figura 16). As frequências destes motivos de interação diferem entre as populações deste estudo, entretanto, essas diferenças são menos pronunciadas que as diferenças observadas para as frequências alélicas (Tabelas 6 e 7).

| Peptídeo Sinal | | D0 | | | | | | | | | | D1 | | | | | | D2 | | | | | | TM-Cit | |
|-----------------------|-------|----|----|----|----|----|----|----|----|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--------|-----|
| Alelo | Códon | 2 | 30 | 31 | 44 | 47 | 54 | 58 | 86 | 92 | 138 | 141 | 145 | 163 | 166 | 182 | 199 | 226 | 238 | 256 | 265 | 283 | 320 | 343 | 373 |
| KIR3DL1*00101 | S L | M | Y | R | R | I | I | S | S | V | G | K | R | P | L | P | P | R | G | Q | T | W | I | C | E |
| KIR3DL1*002 | L F | V | . | . | . | V | L | . | . | . | . | . | . | . | . | . | . | R | . | . | . | . | . | . | . |
| KIR3DL1*00401/02 | . | . | . | . | H | G | . | . | . | L | . | . | . | . | . | S | . | . | . | . | . | L | V | Y | Q |
| KIR3DL1*00501 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | S | . | . | . | . | . | L | . | . | . |
| KIR3DL1*00701 | L F | V | . | . | . | V | L | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | V | . | Q |
| KIR3DL1*008 | L F | V | . | H | . | V | L | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| KIR3DL1*009 | . | . | . | . | . | V | . | G | . | M | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| KIR3DL1*01501/02 | L F | V | . | . | . | V | L | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| KIR3DL1*016 | L | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| KIR3DL1*01701/02 | L F | . | . | . | . | V | L | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| KIR3DL1*018 | . | F | V | . | . | V | L | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| KIR3DL1*019 | . | . | . | C | H | G | . | . | . | L | . | . | . | . | . | S | . | . | . | . | . | L | V | Y | Q |
| KIR3DL1*020 | L F | V | . | . | . | V | L | . | . | . | . | S | . | . | . | . | . | . | . | . | . | . | . | . | . |
| KIR3DL1*025 | L F | V | . | . | . | V | L | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | Q |
| KIR3DL1*02901/v.309*A | L F | V | . | . | . | V | L | . | . | . | . | . | . | . | . | . | . | . | K | . | . | . | . | . | . |
| KIR3DL1*033 | L F | V | . | . | . | V | L | . | . | . | . | E | . | . | . | . | . | . | . | . | . | . | V | . | Q |
| KIR3DL1*041 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | S | . | . | . | . | . | L | V | . | Q |
| KIR3DS1*01301/v.732*C | L | . | . | . | . | V | . | G | . | M | W | . | . | S | R | . | L | . | . | . | . | . | . | . | . |
| v.KIR3DL1*Ex5_740A | L F | V | . | . | . | V | L | . | . | . | . | . | . | . | . | . | . | Q | . | . | . | . | . | . | . |
| v.KIR3DS1*Ex5_857T | L | . | . | . | . | V | . | G | . | M | W | . | . | S | R | . | L | . | . | . | . | I | . | . | . |

FIGURA 15. Sequências de aminoácidos das proteínas codificadas pelos alelos *KIR3DL1/S1* identificados neste estudo. Os pares de alelos *3DL1*00401* e **00402*, *3DL1*01501* e **01502*, *3DL1*01701* e **01702*, *3DL1*02901* e *v.309*A*, *3DS1*01301* e *v.732*C* codificam polipeptídeos com a mesma sequência de aminoácidos. D0, D1 e D2 - Domínios extracelulares; TM-Cit - Domínio transmembrana e citoplasmático.

| Código | N.alelos | D1 | | | | | | | | | | | | | | | | D2 | | | | | | | | | | | | | | | | | | | |
|-------------|----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| | | 115 | 116 | 117 | 118 | 138 | 139 | 140 | 141 | 162 | 163 | 164 | 165 | 166 | 167 | 168 | 169 | 198 | 199 | 200 | 201 | 202 | 203 | 225 | 226 | 227 | 228 | 229 | 230 | 276 | 277 | 278 | 279 | 280 | 281 | 282 | 283 |
| Motivo | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 DL1+015 | 15 | S | G | E | R | G | I | S | K | G | P | M | M | L | A | L | A | G | P | Y | E | K | P | S | R | S | S | Y | D | F | R | H | S | P | Y | E | W |
| 2 DL1+005 | 5 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | L |
| 3 DS1++013 | 3 | . | . | . | W | . | . | . | S | . | . | R | . | . | . | . | . | L | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| 4 DL1+033 | 1 | . | . | . | . | . | E | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| 5 DL1+ 740A | 1 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | Q | . | . | . | . | . | . | . | . | . | . | . | . | . | . |

FIGURA 16. Motivos de interação com as moléculas HLA formados por sítios nos domínios extracelulares D1 e D2 dos receptores KIR3DL1/S1. Motivo 1 - (alelos *3DL1*00101*, **002*, **00701*, **008*, **009*, **01501*, **01502*, **016*, **01701*, **01702*, **018*, **020*, **025*, **02901*, *v.3DL1*309A*); Motivo 2 - (alelos *3DL1*00401*, **00402*, **00501*, **019*, **041*); Motivo 3 - (alelos *3DS1*01301*, *v.3DS1*732C*, *v.3DS1*857T*); Motivo 4 - alelo *3DL1*033*; Motivo 5 - alelo *v.3DL1*740A*.

TABELA 6. Frequência dos motivos de interação com HLA nas diferentes populações.

| | AFRI (2n=46) | EURO-BR (2n=102) | ASI-BR (2n=112) | GKW (2n=128) | GND (2n=116) | GRC (2n=134) | KIV (2n=120) | KRC (2n=144) |
|----------|-----------------|---------------------|--------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Motivo 1 | 34 (73,9) | 46 (45,1) | 80 (71,4) | 70 (54,7) | 63 (54,3) | 62 (46,3) | 53 (44,2) | 62 (43,1) |
| Motivo 2 | 8 (17,4) | 31 (30,4) | 8 (7,1) | 2 (1,6) | 13 (11,2) | 1 (0,7) | 13 (10,8) | 21 (14,6) |
| motivo 3 | 2 (4,3) | 25 (24,5) | 24 (21,4) | 56 (43,8) | 40 (34,5) | 71 (53,0) | 54 (45,0) | 61 (42,4) |
| Motivo 4 | 1 (2,2) | | | | | | | |
| Motivo 5 | 1 (2,2) | | | | | | | |

AFRI - Africanos; EURO-BR - Eurobrasileiros; ASI-BR - brasileiros descendentes de asiáticos; GKW - Guarani Kaiowá; GND - Guarani Nandeva; GRC - Guarani M'byá; KIV - Kaingang de Ivaí; KRC - Kaingang de Rio das Cobras. Entre parênteses são mostradas as frequências relativas para cada motivo, como porcentagens.

TABELA 7. Teste exato de diferenciação populacional baseado nas frequências dos motivos de ligação à HLA.

| | AFRI | EURO-BR | ASI-BR | GKW | GND | GRC | KIV | KRC |
|---------|-------------------|-------------------|-------------------|-------------------|---------|-------------------|---------|-----|
| AFRI | * | | | | | | | |
| EURO-BR | 0,00017 | * | | | | | | |
| ASI-BR | 0,00328 | <10 ⁻⁵ | * | | | | | |
| GKW | <10 ⁻⁵ | <10 ⁻⁵ | <10 ⁻⁴ | * | | | | |
| GND | <10 ⁻⁵ | 0,00206 | 0,02690 | 0,00572 | * | | | |
| GRC | <10 ⁻⁵ | <10 ⁻⁵ | <10 ⁻⁵ | 0,35275 | 0,00004 | * | | |
| KIV | <10 ⁻⁵ | 0,00016 | 0,00001 | 0,00463 | 0,24524 | 0,00064 | * | |
| KRC | <10 ⁻⁵ | 0,00212 | 0,00011 | <10 ⁻⁵ | 0,18582 | <10 ⁻⁵ | 0,68285 | * |

AFRI - Africanos; EURO-BR - Eurobrasileiros; ASI-BR - brasileiros descendentes de asiáticos; GKW - Guarani Kaiowá; GND - Guarani Nandeva; GRC - Guarani M'byá; KIV - Kaingang de Ivaí; KRC - Kaingang de Rio das Cobras.

Os motivos 1 (*DL1*015*) e 2 (*DL1*005*) são comuns a todas as populações e possuem frequências elevadas (Figura 17). Já o motivo 3 (*DS1*013*), embora seja compartilhado por todas as populações, apresenta frequência bastante reduzida na população africana. Outros dois motivos de ligação às moléculas HLA (*DL1*033* e *DL1*v.740A*) foram identificados apenas na população africana.

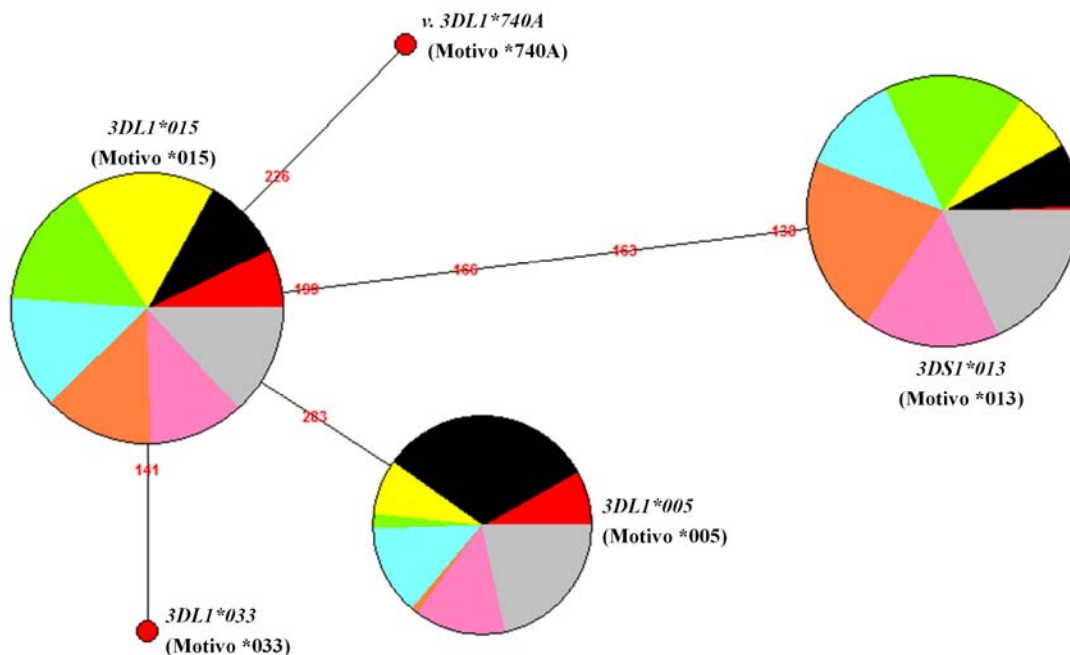


FIGURA 17. Rede mutacional dos motivos de ligação às moléculas HLA dos alelos *KIR3DL1/S1* identificados neste estudo. O tamanho dos discos é proporcional à frequência dos motivos na amostra total deste estudo. Os números em vermelho mostram as posições em que ocorrem diferenças na sequência de aminoácidos. As distâncias não são proporcionais ao número de diferenças. Preto - Eurobrasileiros (EURO-BR); Amarelo - Brasileiros descendentes de asiáticos (ASI-BR); Vermelho - Africanos (AFRI); Verde - GKW; Azul - GND; Salmão - GRC; Rosa - KIV; Cinza - KRC.

5.4 DIVERSIDADE NOS DOMÍNIOS DE INTERAÇÃO COM HLA

Os exons 3, 4 e 5, que codificam os domínios extracelulares que interagem com as moléculas HLA, são as regiões mais variáveis do gene *KIR3DL1/S1* (Figura 18). Esses exons concentram o maior número de substituições nucleotídicas para este gene (Tabela 8) e codificam domínios com diferentes afinidades às moléculas HLA.

FIGURA 18. Regiões variáveis dos exons 3 (D0), 4 (D1) e 5 (D2) dos alelos conhecidos do gene *KIR3DL1/S1*. Em cinza são mostradas as novas variantes identificadas neste estudo.

TABELA 8. Número de substituições sinônimas e não sinônimas nos diferentes domínios do gene *KIR3DL1/S1*.

| Região | Ns | Ms | Ms/Ns | Na | Ma | Ma/Na |
|---|-------|----|-------|-------|----|-------|
| Exons 1 e 2 (Peptídeo sinal) | 14,8 | 0 | 0,000 | 54,2 | 3 | 0,055 |
| Exons 3, 4 e 5 (Domínios extracelulares) | 216,8 | 14 | 0,065 | 663,2 | 44 | 0,066 |
| Exon 7 (Domínio transmembrana) | 25,0 | 5 | 0,200 | 80,0 | 14 | 0,175 |

Ns - Número médio de sítios sinônimos; Ms - Número de substituições sinônimas;
Na - Número médio de sítios não-sinônimos; Ma - Número de substituições não sinônimas.

Devido ao tamanho reduzido das sequências, a proporção de substituições não sinônimas por sítio não sinônimo (Ma/Na) e de substituições sinônimas por sítio sinônimo (Ms/Ns), nos exons 1, 2 e 7, foi verificada através do teste exato de Fisher (tabela 9). As proporções não diferem significativamente em nenhuma das situações consideradas.

Analisando a sequência completa do gene *KIR3DL1/S1* e também apenas exons 3, 4 e 5 do conjunto total de alelos e, ainda, analisando separadamente os alelos *3DL1* e *3DS1*, observa-se que as proporções dN/dS não diferem significativamente para a maioria das situações consideradas (dN=dS) (Tabela 10). Entretanto, o desvio com relação ao valor esperado na situação de neutralidade seletiva é maior se considerada apenas a sequência dos exons 3, 4 e 5 dos alelos *3DS1* ($p=0,059$, para dN=dS). Testado a hipótese de dN ser maior que dS (dN>dS) nos exons 3, 4 e 5 dos alelos *3DS1*, nota-se um acúmulo significativo de substituições não sinônimas ($p=0,0302$). Além disso, analisando apenas os códons que codificam os resíduos que interagem diretamente com as moléculas HLA, nos exons 4 e 5 (códon 115-118, 138-141, 162-169, 198-203, 225-230 e 276-283) do conjunto de alelos *3DL1* e *3DS1* identificados nas populações deste estudo, foi observado um excesso de substituições não sinônimas em relação as substituições sinônimas (dN>dS; $p=0,006$).

TABELA 9. Proporção dN/dS (teste dN=dS) na sequência completa do gene *KIR 3DL1/S1* e nos exons 3, 4 e 5, que codificam os domínios extracelulares.

| Região | dN-dS | p |
|--------------------|--------------|----------|
| Sequência completa | 0,004 | 0,997 |
| Ex3-4e5_3DL1/S1 | -0,870 | 0,386 |
| Ex3-4e5_3DL1 | -0,578 | 0,564 |
| Ex3-4e5_3DS1 | 1,910 | 0,059 |

p - Probabilidade.

Analisando o índice de fixação F_{st} , por sítio polimórfico do conjunto de alelos encontrados neste estudo, observa-se que os menores valores ocorrem nos sítios localizados dentro dos exons que codificam os domínios responsáveis pela interação com HLA (Figura 19). Dentre os sítios específicos que interagem com as moléculas HLA, apenas sete foram polimórficos nas populações deste estudo (códon 138, 141, 163, 166, 199, 226 e 283), sendo que esses sítios apresentam pequena diferença entre as populações estudadas (baixos valores de F_{st}). O valor médio de F_{st} entre todos os sítios segregantes ($F_{st}=0,08403$) confirma a grande diferenciação observada entre as populações analisadas.

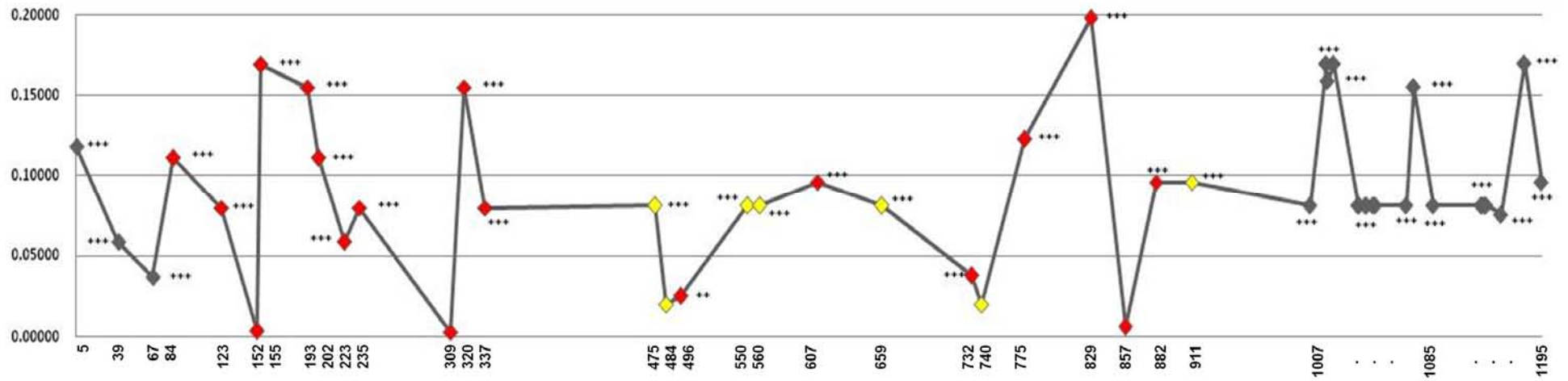


FIGURA 19. Valores globais de F_{st} para os sítios polimórficos encontrados ao longo das sequências dos alelos identificados neste estudo. Em vermelho são mostrados os sítios localizados nos exons que codificam os domínios de interação com HLA. Em amarelo os sítios que fazem parte do sulco de interação. Eixo x: posições polimórficas; Eixo y: valores de F_{st} ; *** $p < 10^{-5}$; ** $p < 0,001$.

As frequências dos ligantes HLA com epítipo Bw4 e Bw6 e das linhagens *3DL1* e *3DS1* (Tabela 10) foram utilizadas para verificar se a probabilidade de coocorrência entre os genótipos de *KIR3DL1/S1* (*3DL1*, *3DS1* e *3DL1/S1*) e os genótipos *HLA* com epítipos *Bw4* e/ou *Bw6* diferem do esperado ao acaso. Ao verificar a distribuição dos epítipos Bw4 e Bw6 entre os indivíduos portadores dos diferentes genótipos *KIR3DL1/S1*, pelo teste exato de Fisher, nenhum desvio significativo foi observado em todas as populações consideradas. Entretanto, o teste de correlação de postos de Spearman revelou uma correlação positiva moderada entre as frequências de *3DL1* e *HLA Bw4* nas populações analisadas (0,671). Ou seja, frequências altas, intermediárias e baixas de *3DL1* coincidem com frequências altas, intermediárias e baixas de *Bw4*. Como esperado devido à correlação negativa entre *3DL1* e *3DS1* e entre *Bw4* e *Bw6*, o teste de correlação de Spearman também revela correlações negativas entre *3DL1* e *Bw6* e entre *3DS1* e *Bw4* e positiva entre *3DS1* e *Bw6*. Uma correlação positiva fraca foi observada entre as frequências de *3DL1* e *HLA Bw4* através do teste de correlação de Pearson.

TABELA 10. Frequências dos epítipos *Bw4* e *Bw6* e alelos *KIR3DL1* e *3DS1* nas populações deste estudo.

| | <i>HLA-Bw4</i> | <i>HLA-Bw6</i> | <i>3DL1</i> | <i>3DS1</i> |
|--------------------|----------------|----------------|-------------|-------------|
| EURO-BR (2n=74) | 21 (28,4) | 53 (71,6) | 61 (82,4) | 13 (17,6) |
| GKW (2n=128) | 14 (10,9) | 114 (89,1) | 71 (56,3) | 56 (43,7) |
| GND (2n=120) | 16 (13,3) | 104 (86,7) | 79 (65,8) | 41 (34,2) |
| GRC (2n=134) | 9 (6,7) | 125 (93,3) | 63 (47,0) | 71 (53,0) |
| KIV (2n=108) | 28 (25,9) | 80 (74,1) | 61 (56,5) | 57 (43,5) |
| KRC (2n=132) | 48 (36,4) | 94 (71,2) | 82 (57,7) | 60 (42,3) |

EURO-BR - Eurodescendentes; GKW: Guarani Kaiowá; GND: Guarani Nandeva; GRC: Guarani M'byá; KIV: Kaingang de Ivaí; KRC: Kaingang de Rio das Cobras. Entre parênteses são mostradas as frequências relativas, como porcentagens. As populações AFRI e ASI-BR não possuem dados de tipagem *HLA*.

5.5 DIVERSIDADE NUCLEOTÍDICA E TESTES DE NEUTRALIDADE SELETIVA

Levando em consideração o tamanho amostral de cada população deste estudo e as frequências alélicas observadas, a diversidade nucleotídica () foi menor na população africana em relação às demais (Tabela 11), porém as médias não diferem significativamente

em nenhuma comparação. Por outro lado, excluindo os alelos *3DS1*, muito frequentes em todas as populações com exceção da população africana, a diversidade nucleotídica desta população é superior a dos indígenas e ASI-BR.

TABELA 11. Diversidade nucleotídica observada nas diferentes populações.

| | AFRI (n=23) | EURO-BR (n=51) | ASI-BR (n=56) | GKW (n=64) | GND (n=58) | GRC (n=67) | KIV (n=60) | KRC (n=72) |
|--------------|----------------|-------------------|------------------|---------------|---------------|---------------|---------------|---------------|
| | 6,984 | 13,351 | 9,466 | 11,932 | 12,501 | 11,402 | 12,213 | 12,416 |
| ^a | 6,038 | 8,116 | 3,642 | 2,313 | 5,005 | 1,290 | 3,731 | 4,100 |

($\times 10^{-4}$) - Diversidade nucleotídica; ^a ($\times 10^{-4}$) - Diversidade nucleotídica considerando apenas os alelos *3DL1*. AFRI - Africanos; EURO-BR - Eurobrasileiros; ASI-BR - brasileiros descendentes de asiáticos; GKW - Guarani Kaiowá; GND - Guarani Nandeva; GRC - Guarani M'byá; KIV - Kaingang de Ivaí; KRC - Kaingang de Rio das Cobras.

Existem vários métodos que permitem verificar se há desvios de neutralidade seletiva sobre sequências genômicas. Um dos testes mais aplicados é o teste D de Tajima. O cálculo deste teste é baseado no número de sítios segregantes na amostra avaliada e no número médio de diferenças nucleotídicas entre cada par de sequências da amostra (TAJIMA, 1989). Valores positivos significativos de D indicam um excesso de variantes (ou alelos) com frequência intermediária, comparado com as frequências esperadas sob neutralidade, e são um indicativo de seleção balanceadora. Por outro lado, valores negativos de D correspondem a um excesso de alelos com baixa frequência e são evidências de seleção purificadora atuando sobre a região estudada. Entretanto valores positivos de D também podem ocorrer devido à subdivisão populacional, enquanto valores negativos podem resultar de uma contração seguida de expansão populacional. Os resultados devem, portanto, ser interpretados com base na história demográfica da população estudada.

Outros testes levam em consideração outros parâmetros. O teste de Homozigosidade de Ewens Watterson (Ewens, 1972; Watterson, 1978) compara a homozigosidade observada com a esperada a partir do número de alelos que ocorrem na população em estudo. Uma homozigosidade observada significativamente superior a esperada pode ser interpretada como seleção direcional e, ao contrário, uma taxa de homozigosidade inferior que a esperada pode indicar seleção balanceadora atuando sobre o loco em estudo.

Já os testes D e F de Fu e LI sem grupo externo (FU e LI, 1993) baseiam-se na diferença entre o número de mutações únicas e o número total de mutações e a média do

número de diferenças nucleotídicas entre pares de sequências, respectivamente. O teste H da Fay e Wu com grupo externo leva em consideração o número médio de diferenças nucleotídicas entre os pares de sequências formados pelo grupo externo e pelas sequências da população em estudo e a frequência dessas variantes na população (FAY e WU, 2000).

Todos os testes de neutralidade não apresentaram significância estatística, embora tenham sido observados valores elevados, como, por exemplo, para o teste D de Tajima (Tabela 12 e 13). Em conjunto, os testes aplicados nas populações deste estudo não permitem rejeitar a hipótese de neutralidade seletiva para o gene *KIR3DL1/S1*. Os testes de neutralidade também foram aplicados considerando apenas os alelos *3DL1*, para verificar se, somente sobre estes alelos, há indícios de seleção natural. Assim como para o conjunto de alelos *3DL1* e *3DS1*, os resultados dos testes não permitem rejeitar a hipótese de neutralidade seletiva (Tabela 13).

TABELA 12. Parâmetros moleculares e testes de neutralidade seletiva para *3DL 1/S1*, nas populações deste estudo.

| Região | $\frac{d_i}{d_e}$ | AFRI | EURO-BR | ASI-BR | GKW | GND | GRC | KIV | KRC |
|--|-------------------|--------|---------|--------|--------|--------|--------|--------|--------|
| Sequência Codificadora Completa | N | 46 | 102 | 112 | 128 | 116 | 134 | 120 | 144 |
| | Alelos | 16 | 14 | 10 | 10 | 12 | 6 | 8 | 6 |
| | S | 40 | 38 | 33 | 31 | 38 | 30 | 30 | 30 |
| | | 6,933 | 13,407 | 9,463 | 11,931 | 12,454 | 11,430 | 12,254 | 12,458 |
| | θ_w | 0,115 | 0,139 | 0,254 | 0,328 | 0,211 | 0,351 | 0,329 | 0,339 |
| | D_T | -0,680 | 2,588 | 1,569 | 3,232 | 2,283 | 3,194 | 3,576 | 3,806 |
| | D_F | 1,458 | 1,502 | 1,369 | 1,299 | 0,346 | 0,170 | 1,966 | 1,620 |
| | F | 0,833 | 2,320 | 1,737 | 2,489 | 1,356 | 1,418 | 3,110 | 3,006 |
| | H | -8,487 | 0,483 | -4,065 | -1,823 | -1,109 | -4,705 | -1,375 | -0,427 |
| | | | | | | | | | |
| 5' UTR e Exons 1 e 2 (Peptídeo sinal) | S | 10 | 8 | 8 | 8 | 8 | 8 | 8 | 8 |
| | | 2,245 | 3,275 | 2,879 | 3,168 | 3,416 | 3,035 | 3,388 | 3,462 |
| | θ_w | 0,115 | 0,139 | 0,254 | 0,328 | 0,211 | 0,351 | 0,329 | 0,339 |
| | D_T | -0,038 | 2,727 | 2,157 | 2,689 | 3,025 | 2,500 | 2,999 | 3,220 |
| | D_F | 0,000 | 1,258 | 1,251 | 1,242 | 1,249 | -0,443 | 1,247 | 1,233 |
| | F | -0,337 | 2,085 | 1,854 | 2,071 | 2,208 | -0,689 | 2,198 | 2,291 |
| | H | -1,271 | 0,523 | -1,247 | 1,315 | -0,915 | -2,000 | -0,921 | -0,633 |
| | | | | | | | | | |
| Exons 3, 4 e 5 (Domínios extracelulares) | S | 20 | 17 | 20 | 16 | 16 | 16 | 20 | 18 |
| | | 3,644 | 5,335 | 6,019 | 4,891 | 5,570 | 5,805 | 6,774 | 4,382 |
| | θ_w | 0,115 | 0,328 | 0,211 | 0,351 | 0,329 | 0,339 | 0,139 | 0,254 |
| | D_T | -0,644 | 1,925 | 1,714 | 1,814 | 2,367 | 2,704 | 2,196 | 0,810 |
| | D_F | 0,956 | 0,606 | 0,362 | -1,128 | 1,620 | 1,061 | 0,844 | 0,703 |
| | F | 0,485 | 1,319 | 1,048 | -0,030 | 2,267 | 2,011 | 1,616 | 0,888 |
| | H | -2,709 | -0,684 | 1,133 | -2,383 | -0,357 | 0,294 | 1,569 | -0,442 |
| Exon 7 (Domínio transmembrana) | S | 12 | 8 | 12 | 8 | 8 | 8 | 12 | 9 |
| | | 0,157 | 3,969 | 3,732 | 4,016 | 3,993 | 3,934 | 4,111 | 2,911 |
| | θ_w | 0,115 | 0,328 | 0,211 | 0,351 | 0,329 | 0,339 | 0,139 | 0,254 |
| | D_T | -1,010 | 3,957 | 1,709 | 4,060 | 3,957 | 3,973 | 2,068 | 1,747 |
| | D_F | 1,477 | 1,242 | -0,468 | 1,239 | 1,247 | 1,234 | 1,467 | 1,312 |
| | F | 0,775 | 2,593 | 0,375 | 2,637 | 2,591 | 2,603 | 2,000 | 1,737 |
| | H | 1,399 | 0,630 | 1,812 | 0,390 | 0,524 | 0,735 | 0,579 | -0,592 |

N - número de cromossomos; S - número de sítios segregantes; θ - ($\times 10^{-4}$) diversidade nucleotídica por sítio; θ_w ($\times 10^{-4}$) - teste de homozigosidade de Ewens Watterson; D_T - teste D de Tajima; D_F - teste D de Fu e Li sem grupo externo; F - teste F de Fu e Li sem grupo externo; H - teste H de Fay e Wu com Chimpanzé como grupo externo (genbank: **NC_006486**). AFRI - Africanos; EURO-BR - Eurobrasileiros; ASI-BR - Brasileiros descendentes de asiáticos; GKW - Guarani Kaiowá; GND - Guarani Nandeva; GRC - Guarani M'byá; KIV - Kaingang de Ivaí; KRC - Kaingang de Rio das Cobras.

TABELA 13. Parâmetros moleculares e testes de neutralidade seletiva para os alelos *3DL1* nas populações deste estudo.

| | $\frac{d}{dL}$ | AFRI | EURO-BR | ASI-BR | GKW | GND | GRC | KIV | KRC |
|--------------------------------|----------------|--------|---------|---------|---------|--------|---------|--------|--------|
| | N | 44 | 77 | 88 | 72 | 76 | 63 | 66 | 83 |
| | Alelos | 15 | 13 | 9 | 8 | 10 | 5 | 7 | 5 |
| | S | 23 | 26 | 17 | 13 | 21 | 12 | 12 | 12 |
| Sequência completa <i>3DL1</i> | | 6,038 | 8,116 | 3,642 | 2,313 | 5,005 | 1,290 | 3,731 | 4,100 |
| | θ_w | 0,124 | 0,138 | 0,337 | 0,452 | 0,240 | 0,319 | 0,417 | 0,479 |
| | D_T | 0,469 | 1,652 | 0,235 | -0,389 | 0,507 | -1,404 | 1,356 | 1,922 |
| | D_F | 0,767 | -0,564 | 0,681 | 0,401 | -1,069 | 0,401 | 1,477 | 0,868 |
| | F | 0,788 | 0,315 | 0,616 | 0,149 | -0,586 | 0,149 | 1,700 | 1,469 |
| | H | -3,038 | 0,206 | -5,497* | -4,839* | -3,458 | -7,193* | -1,891 | -1,117 |

N - número de cromossomos; S - número de sítios segregantes; $\frac{d}{dL}$ - ($\times 10^{-4}$) diversidade nucleotídica por sítio; θ_w ($\times 10^{-4}$) - teste de homozigosidade de Ewens Watterson; D_T - teste D de Tajima; D_F - teste D de Fu e Li sem grupo externo (D^*); F - teste F de Fu e Li sem grupo externo (F^*); H - teste H de Fay e Wu com Chimpanzé como grupo externo (genebank: **NC_006486**). * $p < 0,05$. AFRI - Africanos; EURO-BR - Eurobrasileiros; ASI-BR - Brasileiros descendentes de asiáticos; GKW - Guarani Kaiowá; GND - Guarani Nandeva; GRC - Guarani M'byá; KIV - Kaingang de Ivaí; KRC - Kaingang de Rio das Cobras.

Quando analisadas apenas as regiões dos exons 4 e 5 que codificam os domínios que formam os sulcos de interação com as moléculas HLA, pode-se observar que nos ameríndios, há indícios de seleção balanceadora atuando sobre os domínios codificados pelo exon 4, evidenciados pelos valores elevados de D de Tajima (Tabela 14).

TABELA 14. Teste D de Tajima para as regiões que compreendem os resíduos de interação com as moléculas HLA nos exons 4 e 5.

| | AFRI | EURO-BR | ASI-BR | GKW | GND | GRC | KIV | KRC |
|--------------------------------------|--------|---------|--------|----------|---------|----------|---------|--------|
| Domínios Ex4 (posição 52-216) | -1,549 | 0,953 | 0,899 | 2,907*** | 2,495** | 2,979*** | 2,104** | 1,987* |
| Domínios Ex5 (posição 301-558) | -0,278 | 1,778* | 0,089 | 1,349 | 1,349 | 0,868 | 0,993 | 1,503 |
| Domínios Ex4 e 5 (posição 52-558) | -0,917 | 1,871* | 0,420 | 2,018* | 1,946* | 1,717 | 1,699 | 2,038* |

* $0,05 < p < 0,10$; ** $0,01 < p < 0,05$; *** $p < 0,01$; AFRI - Africanos; EURO-BR - Eurobrasileiros; ASI-BR - Brasileiros descendentes de asiáticos; GKW - Guarani Kaiowá; GND - Guarani Nandeva; GRC - Guarani M'byá; KIV - Kaingang de Ivaí; KRC - Kaingang de Rio das Cobras.

5.6 DISTÂNCIAS GENÉTICAS ENTRE AS POPULAÇÕES

As frequências alélicas do gene *KIR3DL1/S1* foram utilizadas como marcadores para o cálculo da distância genética entre as populações utilizando o método de Cavalli-Sforza e Edwards (1967) (Tabela 15). As maiores distâncias genéticas foram observadas entre a população africana (AFRI) e as cinco populações ameríndias analisadas neste estudo. Embora as frequências alélicas difiram entre todas as populações (ver tabela 4), as populações ameríndias são mais próximas (menores distâncias) à população de descendentes de asiáticos (ASI-BR) e a menor distância genética foi observada entre os indígenas Kaingang de Rio das Cobras (KRC) e os Kaingang de Ivaí (KIV). Os indígenas Guarani (GKW, GND e GRC) também apresentam uma distância genética reduzida entre si.

TABELA 15. Distâncias genéticas entre as populações deste estudo.

| | AFRI | EURO-BR | ASI-BR | GKW | GND | GRC | KIV | KRC |
|---------|----------|----------|----------|----------|----------|----------|----------|-----|
| AFRI | * | | | | | | | |
| EURO-BR | 0,063312 | * | | | | | | |
| ASI-BR | 0,069404 | 0,041176 | * | | | | | |
| GKW | 0,108429 | 0,069430 | 0,045308 | * | | | | |
| GND | 0,099539 | 0,051522 | 0,034035 | 0,011575 | * | | | |
| GRC | 0,090319 | 0,067361 | 0,047906 | 0,016097 | 0,027975 | * | | |
| KIV | 0,078407 | 0,045057 | 0,018010 | 0,036136 | 0,030896 | 0,026061 | * | |
| KRC | 0,097782 | 0,053408 | 0,020585 | 0,029545 | 0,025372 | 0,032245 | 0,006003 | * |

Distâncias genéticas estimadas através do método de Cavalli-Sforza e Edwards (1967). AFRI - Africanos; EURO-BR - Eurobrasileiros; ASI-BR - Brasileiros descendentes de asiáticos; GKW - Guarani Kaiowá; GND - Guarani Nandeva; GRC - Guarani M'byá; KIV - Kaingang de Ivaí; KRC - Kaingang de Rio das Cobras.

A partir das distâncias genéticas, foi construído um dendrograma que representa a relação filogenética entre as populações para as frequências alélicas do gene *KIR3DL1/S1* (Figura 20). Neste dendrograma, os ameríndios (Guarani e Kaingang) formam grupos isolados das demais populações e estão mais próximos dos descendentes de asiáticos. A população africana é a população mais distante das demais. Os eurobrasileiros estão localizados em uma posição intermediária entre as demais populações, distantes dos ameríndios e dos africanos.

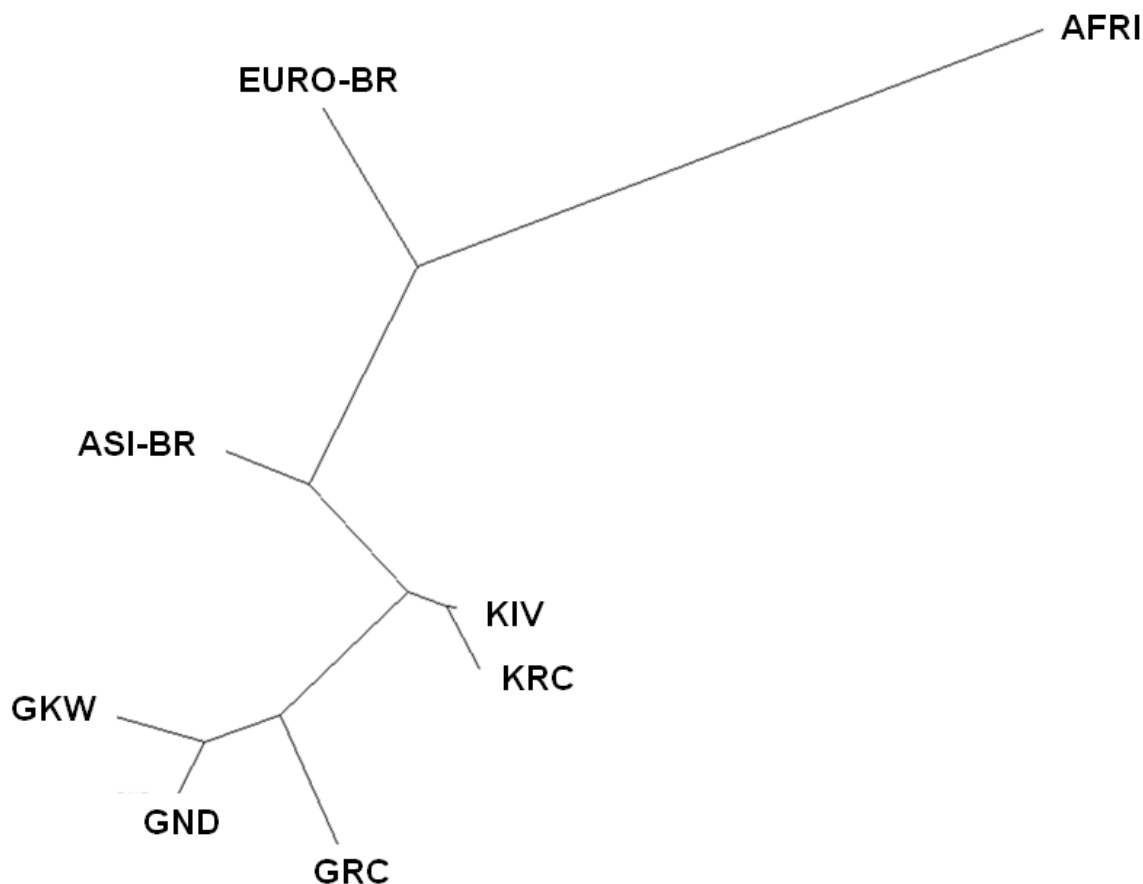


FIGURA 20. Relação entre as populações deste estudo, com base nas frequências alélicas do gene *KIR3DL1/S1*. Método Neighbor-Joining (NJ). AFRI - Africanos; EURO-BR - Eurobrasileiros; ASI-BR - Brasileiros descendentes de asiáticos; GKW - Guarani Kaiowá; GND - Guarani Nandeva; GRC - Guarani M'byá; KIV - Kaingang de Ivaí; KRC - Kaingang de Rio das Cobras.

As distâncias genéticas também foram calculadas entre as populações deste estudo e outras populações para as quais as frequências alélicas do gene *KIR3DL1/S1* estão disponíveis na literatura (Tabela 16). As menores distâncias ocorrem entre os grupos geográfica- e historicamente relacionados, como entre as populações africanas e entre os ameríndios. Estes dois grupos apresentam, entre si, de forma geral, as maiores distâncias genéticas observadas. Entre as populações deste estudo, as distâncias genéticas são menores quando são incluídas outras populações já estudadas, pois, muitos alelos têm frequência igual a zero nas populações deste estudo (estes alelos são observados apenas nas demais populações).

TABELA 16. Distâncias genética entre as populações deste estudo e outras populações.

| População | | | | | | | | | | | | | | | | | | | | | | | | |
|-------------------------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|---|
| AFRI | * | | | | | | | | | | | | | | | | | | | | | | | |
| EURO-BR | 0.0378 | * | | | | | | | | | | | | | | | | | | | | | | |
| ASI-BR | 0.0417 | 0.0248 | * | | | | | | | | | | | | | | | | | | | | | |
| GKW | 0.0647 | 0.0418 | 0.0275 | * | | | | | | | | | | | | | | | | | | | | |
| GND | 0.0578 | 0.0310 | 0.0207 | 0.0065 | * | | | | | | | | | | | | | | | | | | | |
| GRC | 0.0541 | 0.0403 | 0.0287 | 0.0097 | 0.0168 | * | | | | | | | | | | | | | | | | | | |
| KIV | 0.0475 | 0.0269 | 0.0108 | 0.0218 | 0.0186 | 0.0156 | * | | | | | | | | | | | | | | | | | |
| KRC | 0.0587 | 0.0321 | 0.0125 | 0.0181 | 0.0155 | 0.0193 | 0.0036 | * | | | | | | | | | | | | | | | | |
| VEN-BARI ^a | 0.0551 | 0.0348 | 0.0161 | 0.0318 | 0.0313 | 0.0278 | 0.0087 | 0.0067 | * | | | | | | | | | | | | | | | |
| VEN-WARAO ^b | 0.0595 | 0.0343 | 0.0177 | 0.0310 | 0.0244 | 0.0294 | 0.0095 | 0.0053 | 0.0108 | * | | | | | | | | | | | | | | |
| VEN-YUCPA ^c | 0.0668 | 0.0428 | 0.0233 | 0.0070 | 0.0121 | 0.0095 | 0.0122 | 0.0068 | 0.0170 | 0.0159 | * | | | | | | | | | | | | | |
| ORI-SOUTH KOREA ^d | 0.0532 | 0.0308 | 0.0041 | 0.0279 | 0.0202 | 0.0325 | 0.0113 | 0.0082 | 0.0118 | 0.0124 | 0.0196 | * | | | | | | | | | | | | |
| ORI-TAIWAN ^e | 0.0442 | 0.0192 | 0.0050 | 0.0323 | 0.0209 | 0.0395 | 0.0156 | 0.0139 | 0.0180 | 0.0154 | 0.0277 | 0.0064 | * | | | | | | | | | | | |
| ORI-JAPAN ^f | 0.0485 | 0.0324 | 0.0043 | 0.0415 | 0.0298 | 0.0445 | 0.0151 | 0.0147 | 0.0172 | 0.0165 | 0.0321 | 0.0046 | 0.0047 | * | | | | | | | | | | |
| INDIA SOURSHTRAN ^g | 0.0528 | 0.0165 | 0.0195 | 0.0318 | 0.0231 | 0.0419 | 0.0254 | 0.0245 | 0.0252 | 0.0281 | 0.0344 | 0.0216 | 0.0095 | 0.0219 | * | | | | | | | | | |
| INDIA MADURAI ^h | 0.0532 | 0.0135 | 0.0198 | 0.0363 | 0.0268 | 0.0483 | 0.0302 | 0.0291 | 0.0282 | 0.0333 | 0.0405 | 0.0222 | 0.0108 | 0.0233 | 0.0025 | * | | | | | | | | |
| EURO-UK ⁱ | 0.0483 | 0.0072 | 0.0307 | 0.0464 | 0.0346 | 0.0555 | 0.0378 | 0.0382 | 0.0424 | 0.0403 | 0.0486 | 0.0348 | 0.0183 | 0.0340 | 0.0141 | 0.0111 | * | | | | | | | |
| EURO-IRELAND ^j | 0.0533 | 0.0066 | 0.0278 | 0.0453 | 0.0317 | 0.0552 | 0.0382 | 0.0367 | 0.0395 | 0.0380 | 0.0475 | 0.0297 | 0.0158 | 0.0308 | 0.0112 | 0.0069 | 0.0035 | * | | | | | | |
| EURO-GEORGIA ^k | 0.0512 | 0.0083 | 0.0246 | 0.0419 | 0.0316 | 0.0530 | 0.0359 | 0.0363 | 0.0384 | 0.0403 | 0.0466 | 0.0293 | 0.0145 | 0.0290 | 0.0081 | 0.0043 | 0.0037 | 0.0025 | * | | | | | |
| EURO-TURKEY ^l | 0.0471 | 0.0103 | 0.0288 | 0.0424 | 0.0340 | 0.0523 | 0.0355 | 0.0373 | 0.0401 | 0.0426 | 0.0468 | 0.0346 | 0.0186 | 0.0334 | 0.0104 | 0.0090 | 0.0047 | 0.0079 | 0.0029 | * | | | | |
| EURO-SPAIN ^m | 0.0403 | 0.0094 | 0.0242 | 0.0348 | 0.0253 | 0.0422 | 0.0313 | 0.0308 | 0.0356 | 0.0338 | 0.0368 | 0.0286 | 0.0160 | 0.0308 | 0.0131 | 0.0115 | 0.0086 | 0.0080 | 0.0073 | 0.0094 | * | | | |
| AFRI-NIGERIA ⁿ | 0.0182 | 0.0324 | 0.0375 | 0.0595 | 0.0484 | 0.0536 | 0.0429 | 0.0534 | 0.0548 | 0.0565 | 0.0646 | 0.0467 | 0.0397 | 0.0426 | 0.0444 | 0.0438 | 0.0423 | 0.0441 | 0.0423 | 0.0407 | 0.0357 | * | | |
| AFRI-GHANA ^o | 0.0239 | 0.0358 | 0.0402 | 0.0639 | 0.0515 | 0.0617 | 0.0514 | 0.0611 | 0.0612 | 0.0647 | 0.0726 | 0.0492 | 0.0425 | 0.0451 | 0.0452 | 0.0427 | 0.0454 | 0.0441 | 0.0426 | 0.0438 | 0.0388 | 0.0033 | * | |
| AFRI-ZIMBABWE ^p | 0.0180 | 0.0336 | 0.0384 | 0.0556 | 0.0447 | 0.0531 | 0.0444 | 0.0515 | 0.0519 | 0.0529 | 0.0608 | 0.0447 | 0.0368 | 0.0426 | 0.0411 | 0.0408 | 0.0424 | 0.0419 | 0.0421 | 0.0436 | 0.0386 | 0.0078 | 0.0093 | * |

Distâncias genéticas estimadas através do método de Cavalli-Sforza e Edwards (1967). Em cinza as populações analisadas neste estudo. AFRI - Africanos; EURO-BR - Eurobrasileiros; ASI-BR - Brasileiros descendentes de asiáticos; GKW - Guarani Kaiowá; GND - Guarani Nandeva; GRC - Guarani M'byá; KIV - Kaingang de Ivai; KRC - Kaingang de Rio das Cobras. a,b,c - Ameríndios da Venezuela (NORMAN et al., 2007); d - Asiáticos da Coreia do Sul (WHANG et al., 2005); e - Asiáticos de Taiwan (NORMAN et al., 2007); f - Asiáticos Japoneses (YAWATA et al., 2006); g - Asiáticos Sourshtran da Índia (NORMAN et al., 2007); h - Asiáticos Madurai da Índia (NORMAN et al., 2007); i - Europeus do Reino Unido (NORMAN et al., 2007); j - Europeus da Irlanda do Norte (MIDDLETON, 2003); k - Europeus da Georgia (NORMAN et al., 2007); l - Europeus da Turquia (NORMAN et al., 2007); m - Europeus Majorca da Espanha (NORMAN et al., 2007); n - Africanos da Nigéria (NORMAN et al., 2007); o - Africanos de Ghana (NORMAN et al., 2007); p - Africanos do Zimbabwe (NORMAN et al., 2007).

O dendrograma construído a partir das distâncias genéticas entre todas as populações mostra que estas populações se agrupam, de forma geral, conforme sua origem e história evolutiva, formando quatro principais grupos (ameríndios, populações do leste asiático, africanos e eurodescendentes) (Figura 21).

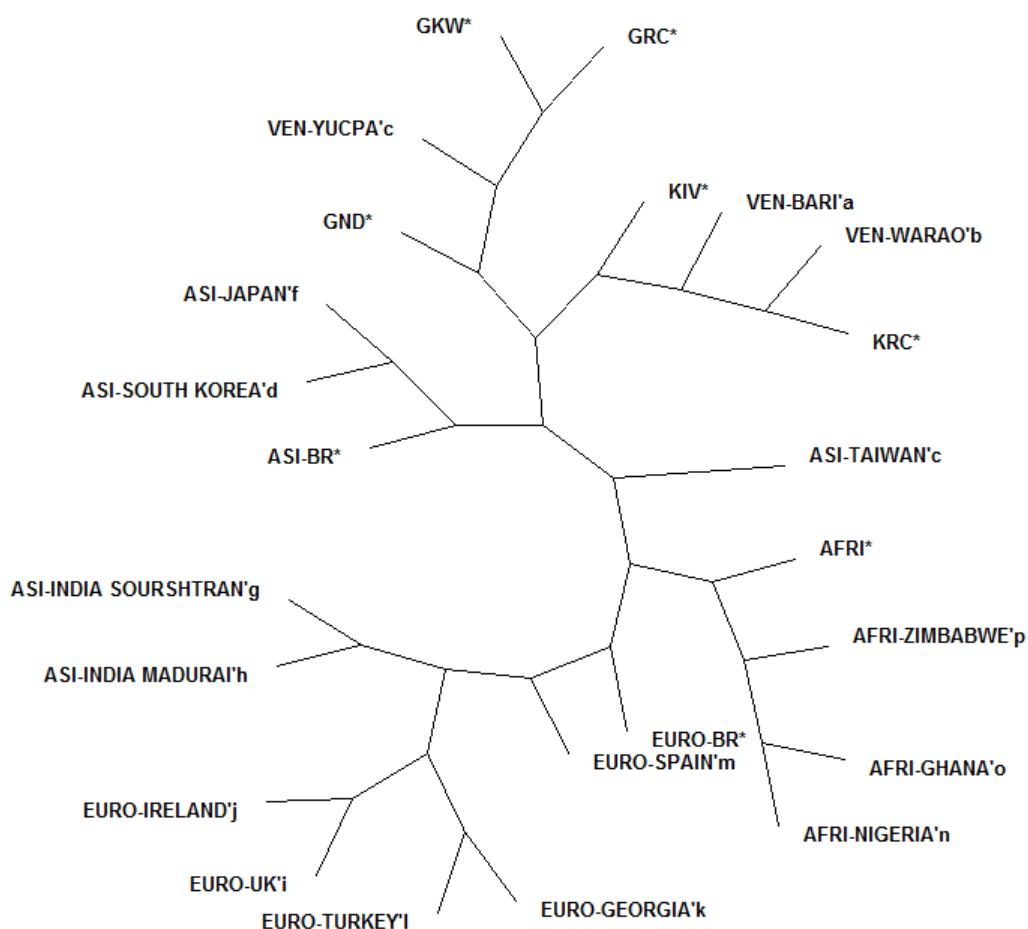


FIGURA 21. Relação entre as populações deste estudo (*) e outras populações da literatura, com base nas frequências alélicas do gene *KIR3DL1/S1*. Método Neighbor-Joining (NJ). AFRI - Africanos; EURO-BR - Eurobrasileiros; ASI-BR - Brasileiros descendentes de asiáticos; GKW - Guarani Kaiowá; GND - Guarani Nandeva; GRC - Guarani M'byá; KIV - Kaingang de Ivaí; KRC - Kaingang de Rio das Cobras. a,b,c - Ameríndios da Venezuela (NORMAN et al., 2007); d - Asiáticos da Coreia do Sul (WHANG et al., 2005); e - Asiáticos de Taiwan (NORMAN et al., 2007); f - Asiáticos Japoneses (YAWATA et al., 2006); g - Asiáticos Sourshtran da Índia (NORMAN et al., 2007); h - Asiáticos Madurai da Índia (NORMAN et al., 2007); i - Europeus do Reino Unido (NORMAN et al., 2007); j - Europeus da Irlanda do Norte (<<http://www.allele frequencies.net>>, MIDDLETON, 2003); k - Europeus da Georgia (NORMAN et al., 2007); l - Europeus da Turquia (NORMAN et al., 2007); m - Europeus Majorca da Espanha (NORMAN et al., 2007); n - Africanos da Nigéria (NORMAN et al., 2007); o - Africanos de Ghana (NORMAN et al., 2007); p - Africanos do Zimbabwe (NORMAN et al., 2007).

Um dendrograma também foi construído levando em consideração as frequências dos motivos de ligação às moléculas HLA nas populações deste estudo (Figura 22). A topologia deste dendrograma difere da observada para as frequências alélicas, porém os ameríndios continuam mais relacionados entre si que as demais populações.

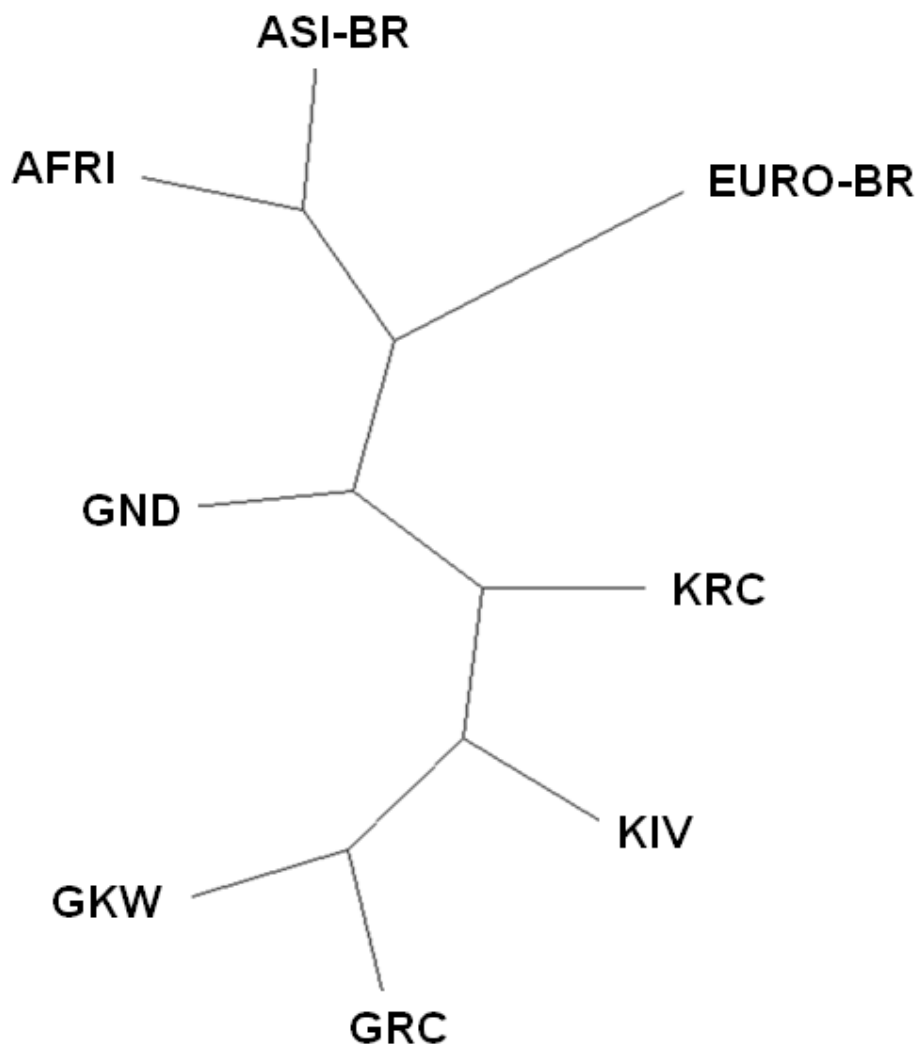


FIGURA 22. Relação entre as populações deste estudo, com base nas frequências dos motivos de ligação às moléculas HLA dos alelos *KIR3DL1/S1* identificados nestas populações. Método Neighbor-Joining (NJ). AFRI - Africanos; EURO-BR - Eurobrasileiros; ASI-BR - Brasileiros descendentes de asiáticos; GKW - Guarani Kaiowá; GND - Guarani Nandeva; GRC - Guarani M'byá; KIV - Kaingang de Ivaí; KRC - Kaingang de Rio das Cobras.

5.7 DESEQUILÍBRIO DE LIGAÇÃO

O desequilíbrio de ligação foi verificado nas populações deste estudo excluindo as deleções e considerando apenas os sítios polimórficos compartilhados entre os alelos *3DL1* e *3DS1* (43 sítios) (Figuras 23 - 30). De modo geral, o desequilíbrio é o máximo possível entre todas as posições e não são observados blocos separados de desequilíbrio. Porém, alguns sítios polimórficos são compartilhados por diferentes alelos (que diferem em outras regiões da sequência), apresentando um menor valor para o desequilíbrio de ligação (por exemplo, SNPs 1 e 2 na população africana). O desequilíbrio foi absoluto para todas as posições nas populações ASI-BR, GRC, KRC e KIV. Devido aos vários alelos encontrados nas populações e as diferentes frequências destes alelos, o valor de r^2 é baixo para a maioria dos SNPs e não é possível identificar apenas um tag SNP que forneça informação para toda a região do gene *KIR3DL1/S1*. Entretanto, na extremidade 3' estão agrupados vários SNPs com $r^2 > 0,9$, em todas as populações. Os ameríndios e descendentes de asiáticos (ASI-BR) apresentam um maior número de posições monomórficas quando comparados com as populações eurobrasileira e africana. Pode-se verificar também que o desequilíbrio de ligação é menor na amostra populacional africana e eurodescendente, em relação às demais, com menores valores de D' e r^2 .

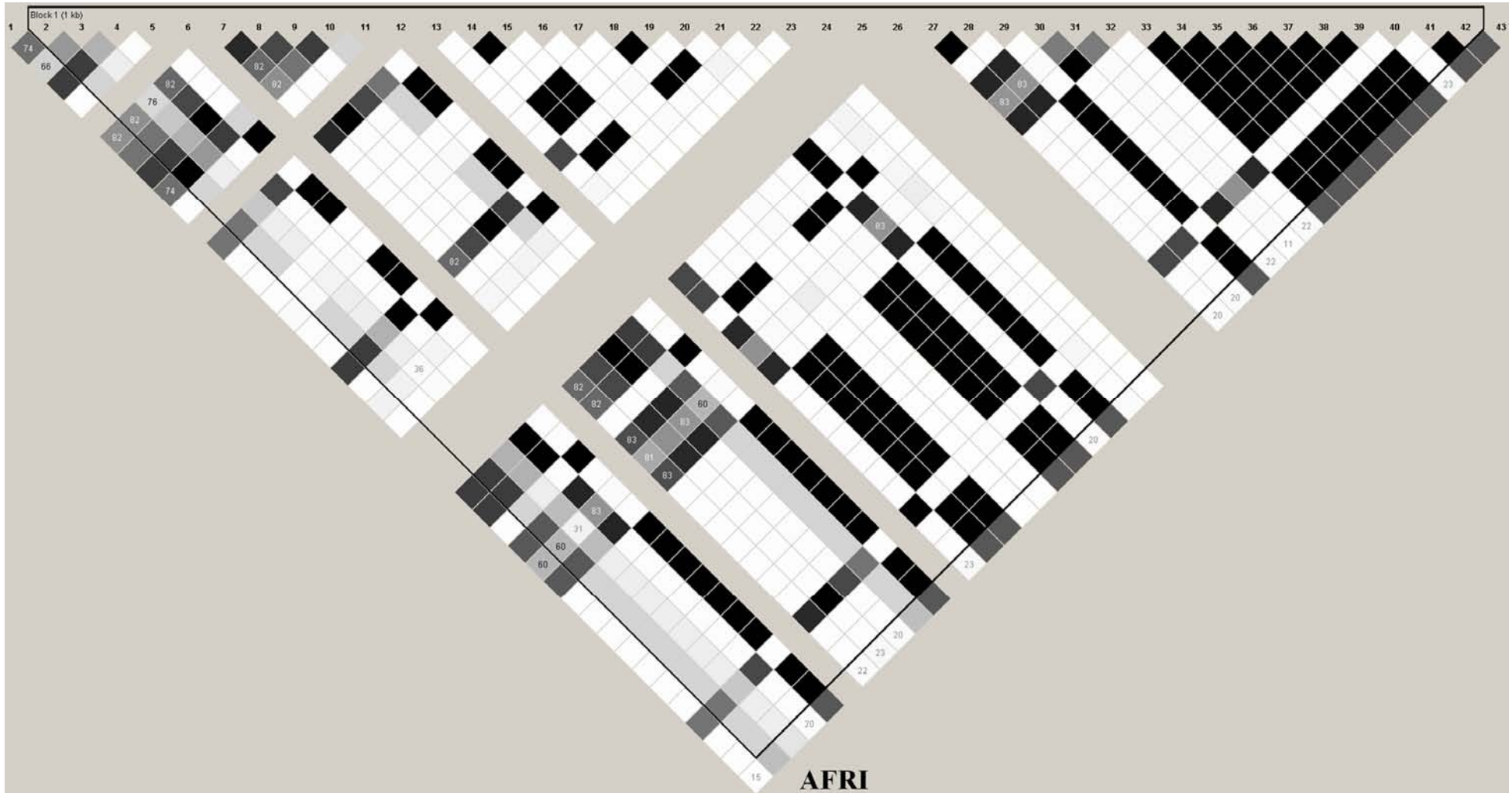


FIGURA 23. Padrão do desequilíbrio de ligação entre os 43 SNPs ao longo da sequência dos alelos identificados neste estudo na população africana (AFRI). Cada losango representa o desequilíbrio de ligação entre dois sítios. Números no interior dos losangos correspondem aos valores de D' . Para todos os losangos vazios, o DL foi o máximo possível ($D'=1$). Losangos pretos representam valores significativos para o desequilíbrio de ligação ($r^2 > 0,9$). Losangos cinza escuro representam valores de $r^2 = 0,7-0,9$; cinza médio, $r^2 = 0,4-0,7$; cinza claro, $r^2 = 0,1-0,4$; Losangos brancos, $r^2 = 0$. Linhas vazias representam posições monomórficas nesta população.

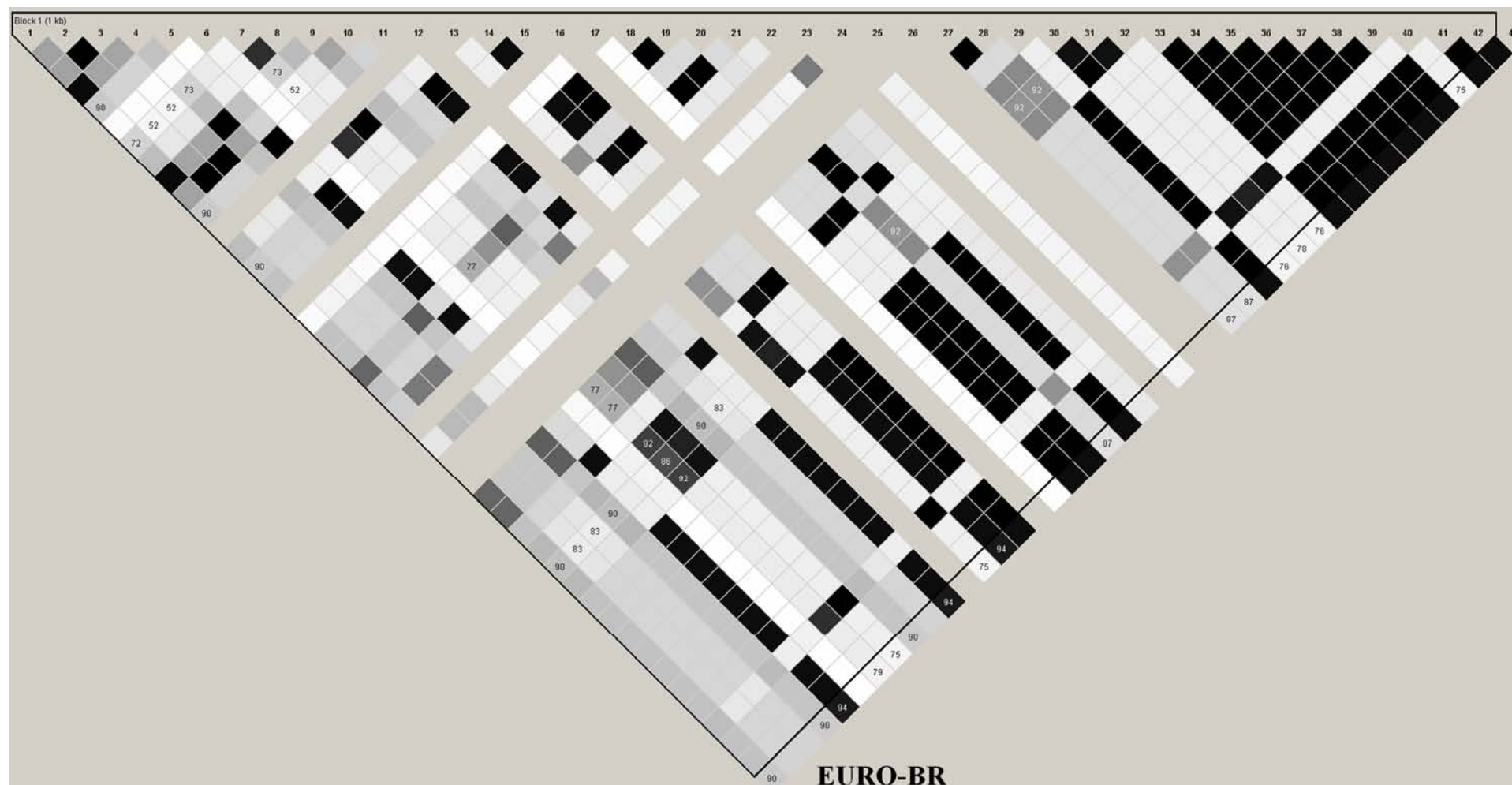


FIGURA 24. Padrão do desequilíbrio de ligação entre os 43 SNPs ao longo da sequência dos alelos identificados neste estudo na população eurobrasileira (EURO-BR). Cada losango representa o desequilíbrio de ligação entre dois sítios. Números no interior dos losangos correspondem aos valores de D' . Para todos os losangos vazios, o DL foi o máximo possível ($D'=1$). Losangos pretos representam valores significativos para o desequilíbrio de ligação ($r^2 > 0,9$). Losangos cinza escuro representam valores de $r^2 = 0,7-0,9$; cinza médio, $r^2 = 0,4-0,7$; cinza claro, $r^2 = 0,1-0,4$; Losangos brancos, $r^2 = 0$. Linhas vazias representam posições monomórficas nesta população.

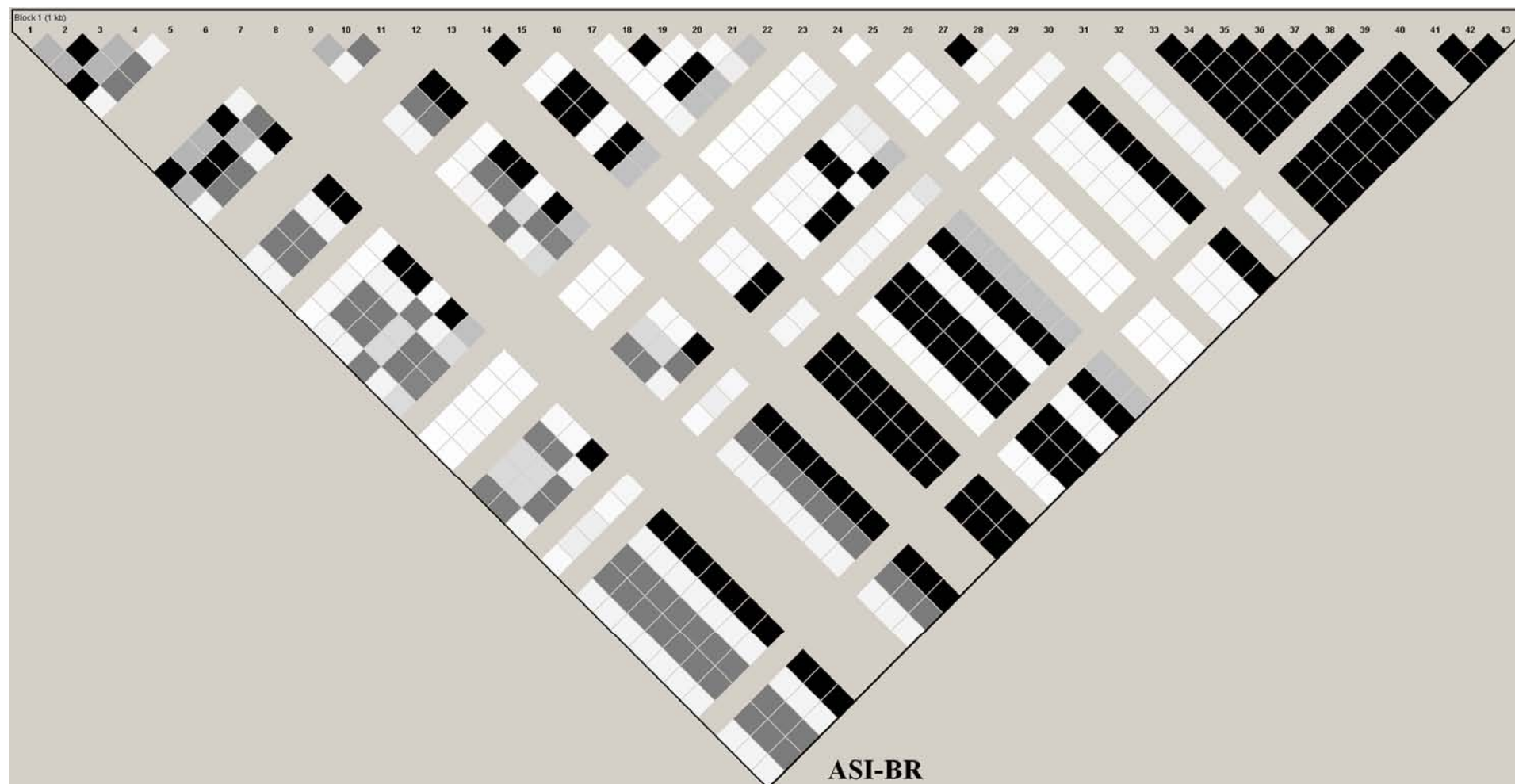


FIGURA 25. Padrão do desequilíbrio de ligação entre os 43 SNPs ao longo da sequência dos alelos identificados neste estudo na população descendente de asiáticos (ASI-BR). Cada losango representa o desequilíbrio de ligação entre dois sítios. Para todos os pares de SNPs o DL foi o máximo possível ($D'=1$). Losangos pretos representam valores significativos para o desequilíbrio de ligação ($r^2 > 0,9$). Losangos cinza escuro representam valores de $r^2 = 0,7-0,9$; cinza médio, $r^2 = 0,4-0,7$; cinza claro, $r^2 = 0,1-0,4$; Losangos brancos, $r^2 = 0$. Linhas vazias representam posições monomórficas nesta população.

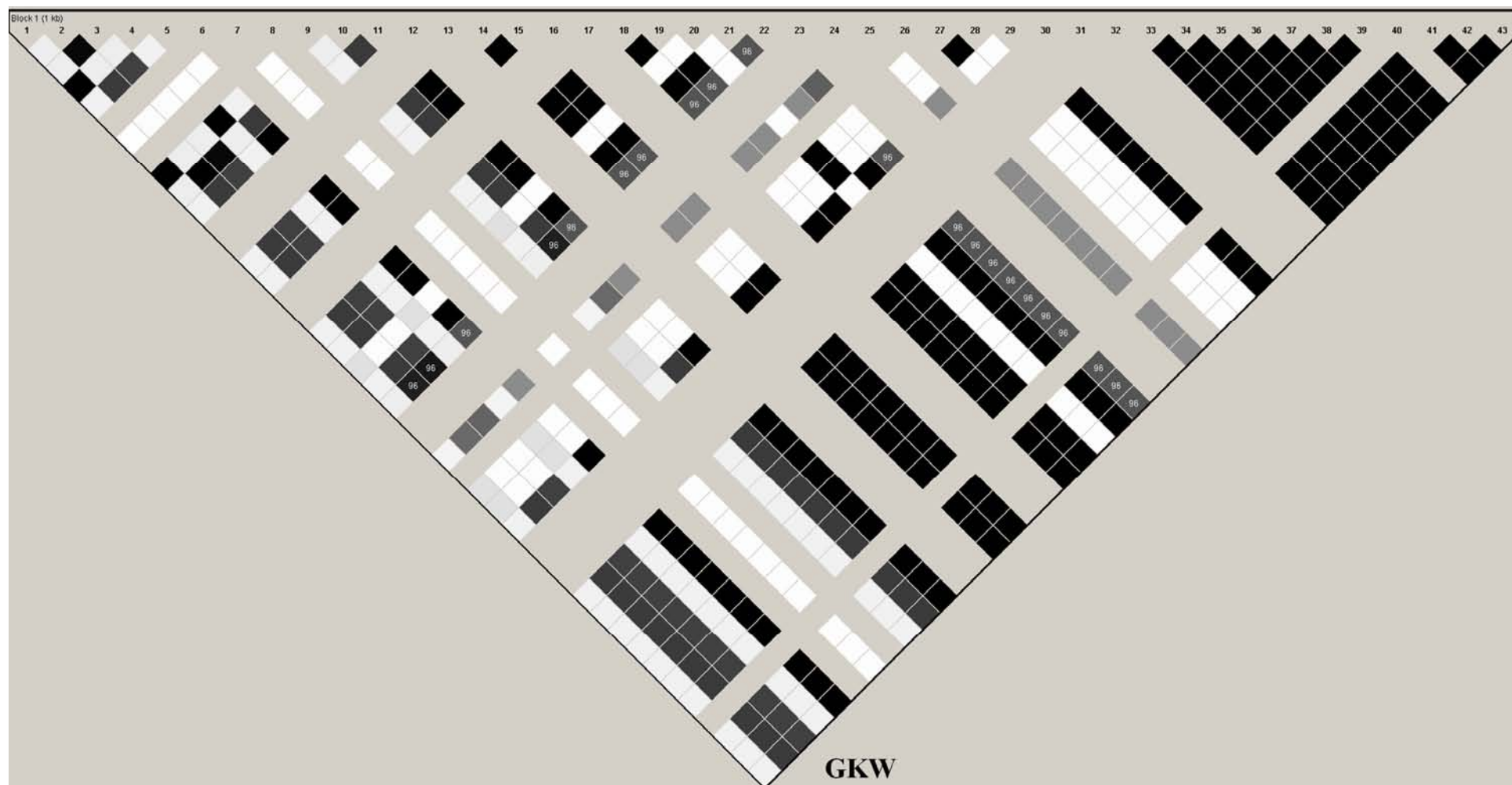


FIGURA 26. Padrão do desequilíbrio de ligação entre os 43 SNPs ao longo da sequência dos alelos identificados neste estudo na população GKW. Cada losango representa o desequilíbrio de ligação entre dois sítios. Números no interior dos losangos correspondem aos valores de D' . Para todos os losangos vazios, o DL foi o máximo possível ($D'=1$). Losangos pretos representam valores significativos para o desequilíbrio de ligação ($r^2 > 0,9$). Losangos cinza escuro representam valores de $r^2 = 0,7-0,9$; cinza médio, $r^2 = 0,4-0,7$; cinza claro, $r^2 = 0,1-0,4$; Losangos brancos, $r^2 = 0$. Linhas vazias representam posições monomórficas nesta população.

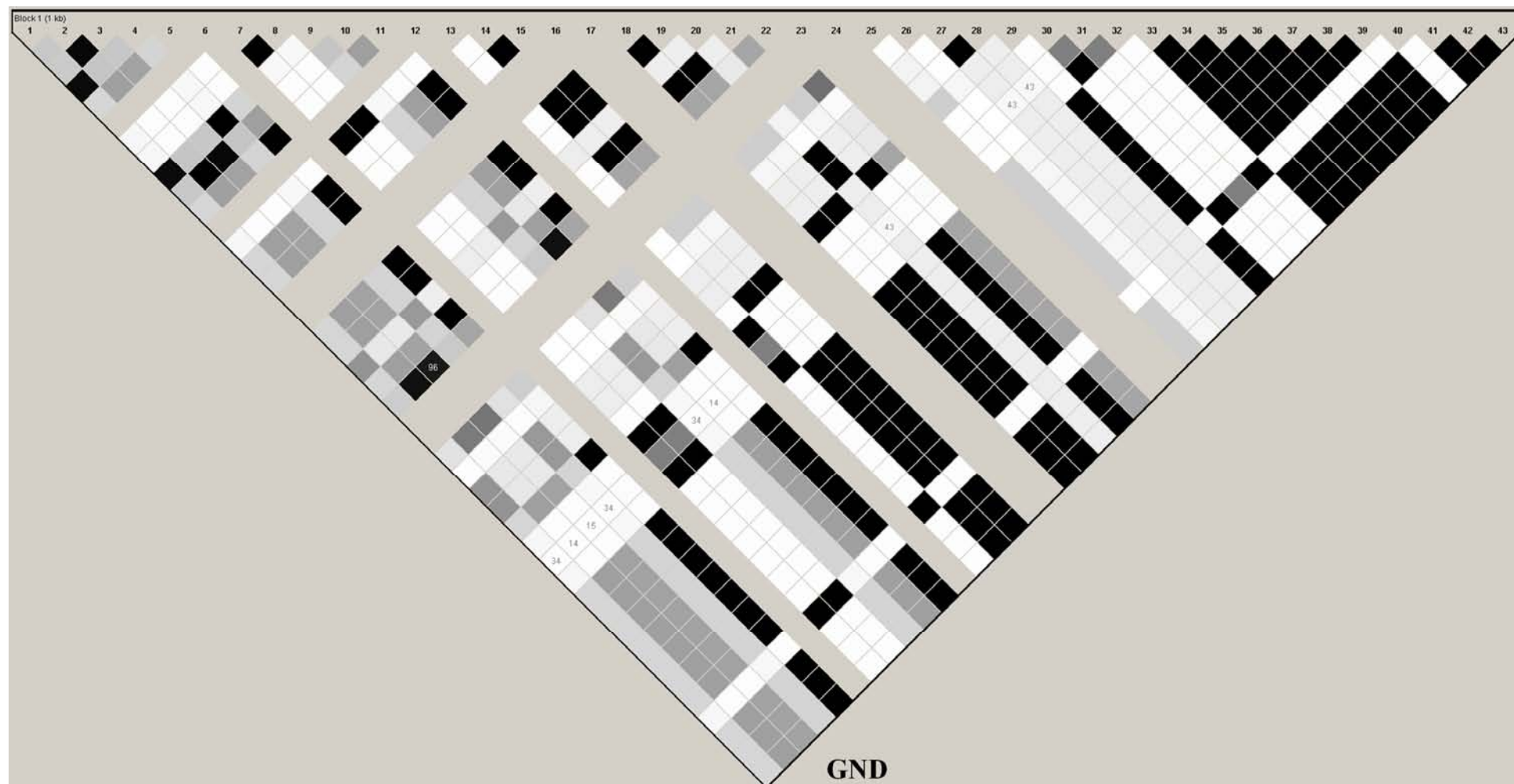


FIGURA 27. Padrão do desequilíbrio de ligação entre os 43 SNPs ao longo da sequência dos alelos identificados neste estudo na população GND. Cada losango representa o desequilíbrio de ligação entre dois sítios. Números no interior dos losangos correspondem aos valores de D' . Para todos os losangos vazios, o DL foi o máximo possível ($D'=1$). Losangos pretos representam valores significativos para o desequilíbrio de ligação ($r^2 > 0,9$). Losangos cinza escuro representam valores de $r^2 = 0,7-0,9$; cinza médio, $r^2 = 0,4-0,7$; cinza claro, $r^2 = 0,1-0,4$; Losangos brancos, $r^2 = 0$. Linhas vazias representam posições monomórficas nesta população.

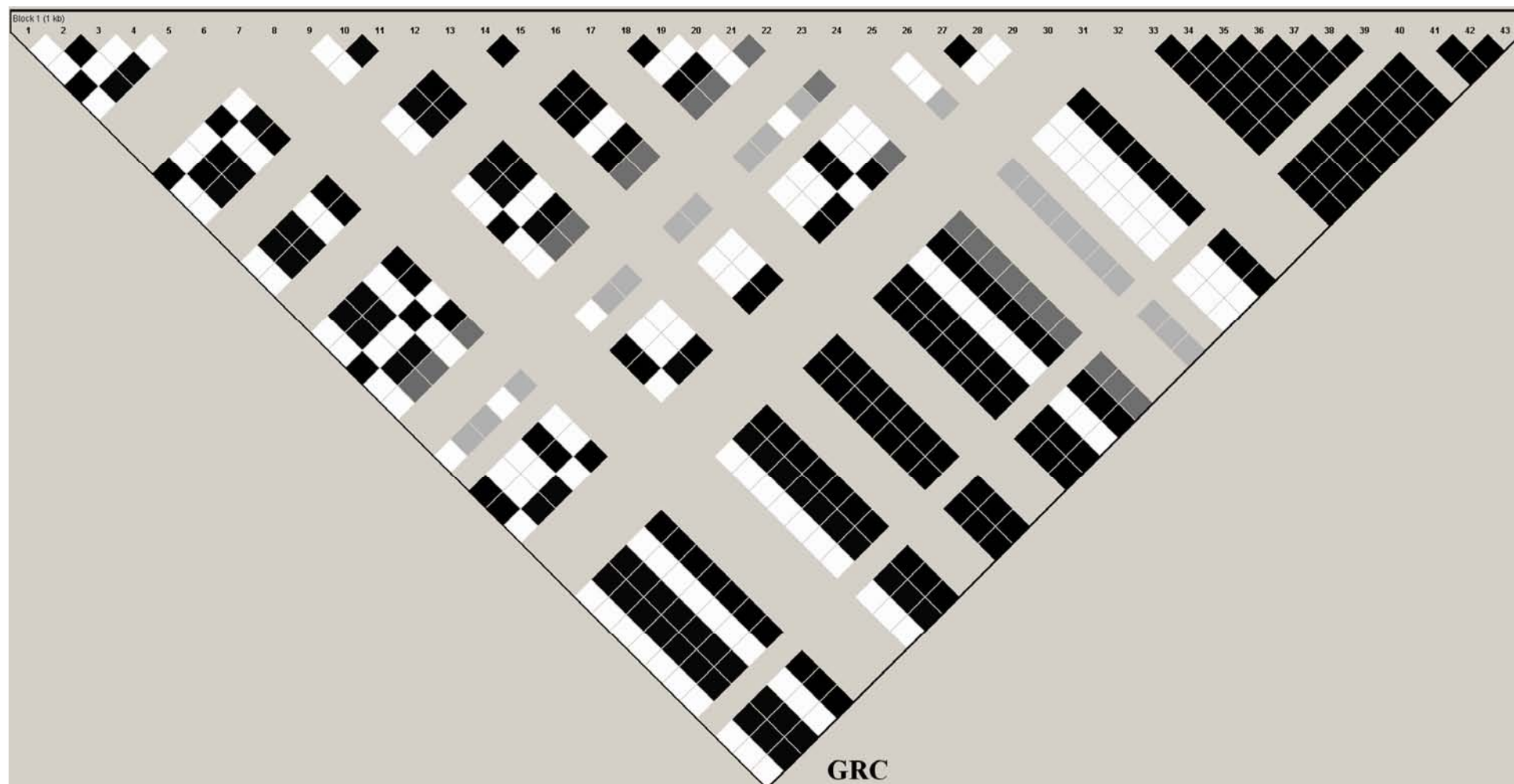


FIGURA 28. Padrão do desequilíbrio de ligação entre os 43 SNPs ao longo da sequência dos alelos identificados neste estudo na população GRC. Cada losango representa o desequilíbrio de ligação entre dois sítios. Para todos os pares de SNPs o DL foi o máximo possível ($D'=1$). Losangos pretos representam valores significativos para o desequilíbrio de ligação ($r^2 > 0,9$). Losangos cinza escuro representam valores de $r^2 = 0,7-0,9$; cinza médio, $r^2 = 0,4-0,7$; cinza claro, $r^2 = 0,1-0,4$; Losangos brancos, $r^2 = 0$. Linhas vazias representam posições monomórficas nesta população.

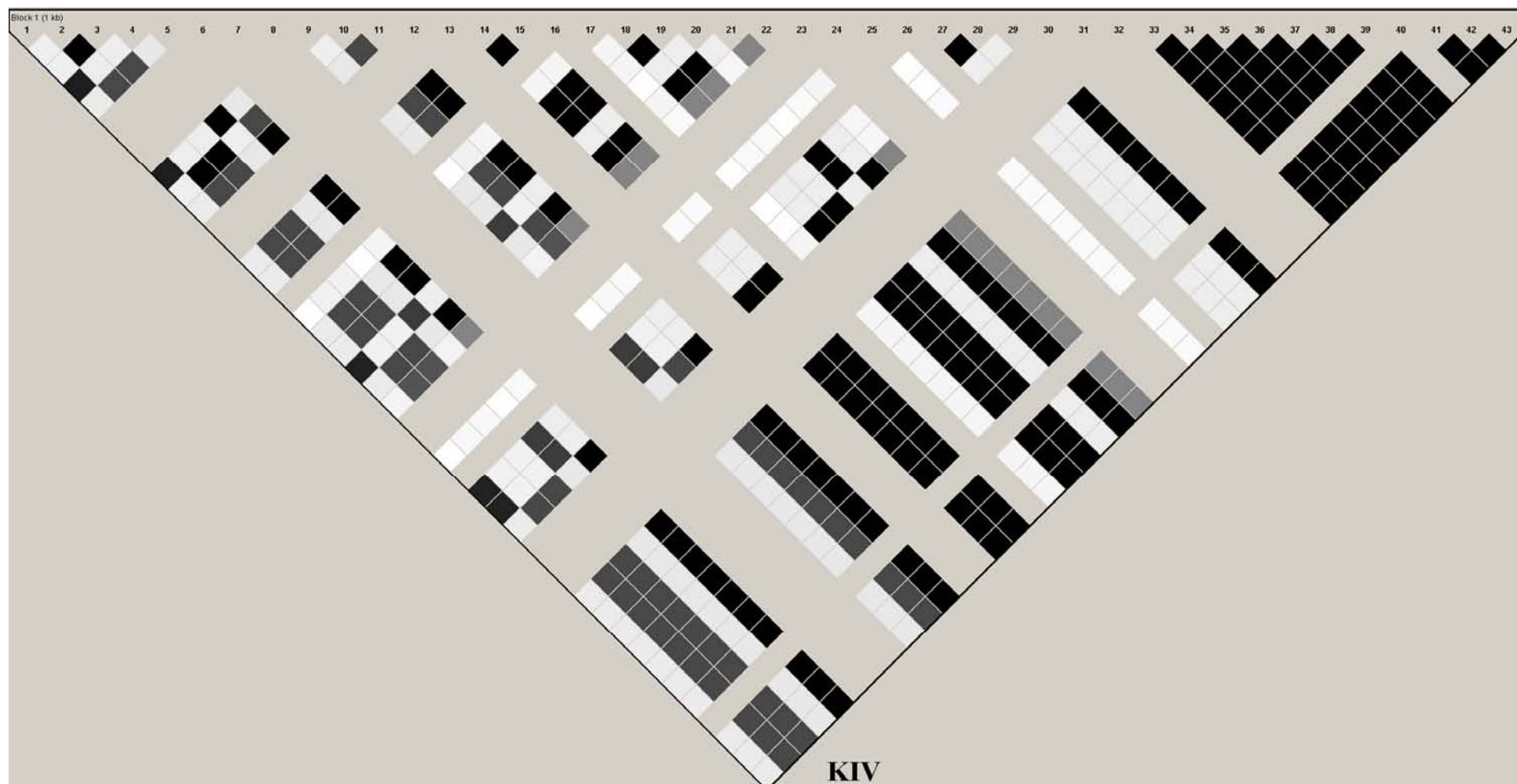


FIGURA 29. Padrão do desequilíbrio de ligação entre os 43 SNPs ao longo da sequência dos alelos identificados neste estudo na população KIV. Cada losango representa o desequilíbrio de ligação entre dois sítios. Para todos os pares de SNPs o DL foi o máximo possível ($D'=1$). Losangos pretos representam valores significativos para o desequilíbrio de ligação ($r^2 > 0,9$). Losangos cinza escuro representam valores de $r^2 = 0,7-0,9$; cinza médio, $r^2 = 0,4-0,7$; cinza claro, $r^2 = 0,1-0,4$; Losangos brancos, $r^2 = 0$. Linhas vazias representam posições monomórficas nesta população.

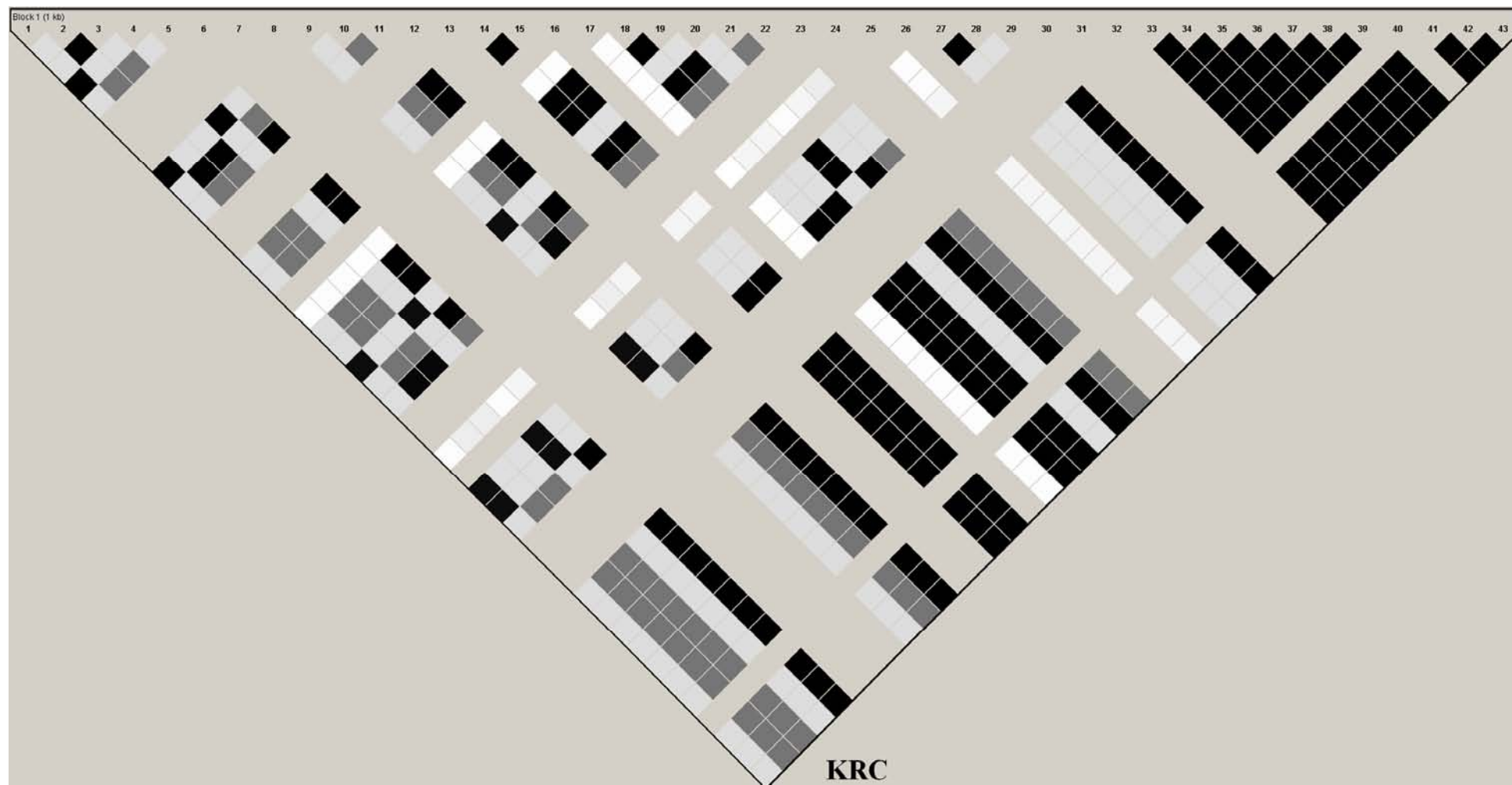


FIGURA 30. Padrão do desequilíbrio de ligação entre os 43 SNPs ao longo da sequência dos alelos identificados neste estudo na população KRC. Cada losango representa o desequilíbrio de ligação entre dois sítios. Para todos os pares de SNPs o DL foi o máximo possível ($D'=1$). Losangos pretos representam valores significativos para o desequilíbrio de ligação ($r^2 > 0,9$). Losangos cinza escuro representam valores de $r^2 = 0,7-0,9$; cinza médio, $r^2 = 0,4-0,7$; cinza claro, $r^2 = 0,1-0,4$; Losangos brancos, $r^2 = 0$. Linhas vazias representam posições monomórficas nesta população.

5.8 ANÁLISE EVOLUTIVA DO GENE *KIR3DL1/S1*

A relação filogenética entre os alelos demonstra que os alelos *KIR3DL1/S1* são bastante diversos entre si e não formam grupos muito concisos (baixos valores de bootstrap) (Figura 31). Apesar disso, três grupos principais, correspondentes às três linhagens ancestrais deste gene, podem ser identificados (linhagem *015, linhagem *005 e linhagem 3DS1*013). Dentre esses grupos, a linhagem 3DS1*013 é sustentada pelo maior valor de bootstrap (67). Os novos alelos 3DL1 identificadas neste estudo (v.3DL1*309A e v.3DL1*740A) pertencem à linhagem *015 e os alelos v.3DS1*732C e v.3DS1*857T são agrupados com os alelos da linhagem 3DS1*013. Observa-se ainda que os alelos humanos são bastante distantes da sequência homóloga *KIR3DL1/2*, encontrada nos chimpanzés (*Pan troglodytes*) (grupo externo neste dendrograma).

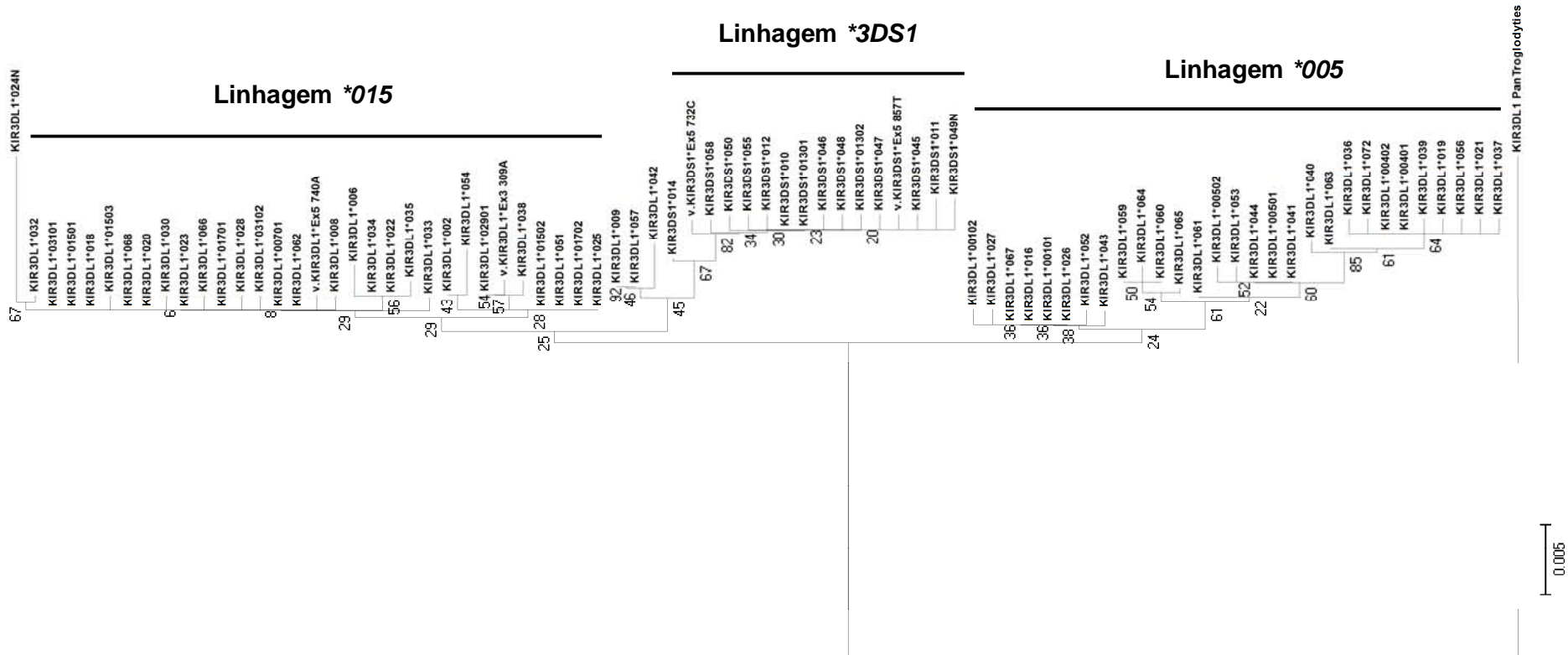


FIGURA 31. Relação filogenética sugerida para os alelos *KIR3DL1/S1* e os novos alelos identificados neste estudo. Método Neighbor-Joining (NJ). Os valores de bootstrap são mostrados nos “nós” do dendrograma.

O agrupamento dos alelos do gene *KIR3DL1/S1* em três linhagens também pode ser observada na rede mutacional dos alelos encontrados nas populações deste estudo (figura 13). Nesta rede, observa-se que os alelos das três linhagens estão mais proximamente relacionados e diferem por um menor número de mudanças nucleotídicas.

6 DISCUSSÃO

Embora a maior parte da variação genética seja encontrada dentro das populações, as frequências alélicas variam entre diferentes populações. Essas diferenças ocorrem devido à ação de fatores evolutivos e/ou efeitos demográficos. A caracterização genética das populações humanas é importante para a reconstrução da história da fundação ou migração e investigação da estrutura genética das populações, possibilitando inferir as relações históricas e evolutivas entre elas. O estudo da diversidade de genes do sistema imune é particularmente interessante, pois estes genes aparentemente estão sujeitos a seleção natural acentuada, especialmente seleção positiva, flutuante e balanceadora, resultando em elevada taxa evolutiva, em comparação com a maioria dos genes de funções não relacionadas às respostas imunes e ao sucesso reprodutivo. O conhecimento de genética de populações é igualmente importante para a medicina evolutiva e para estimativa dos impactos populacionais das associações entre polimorfismos genéticos e doenças. Esse conhecimento pode, por exemplo, ser utilizado para ações de interesse na área de saúde pública.

6.1 DIVERSIDADE GENÉTICA DO LOCO *KIR3DL1/S1*

O loco *KIR3DL1/S1* exibe uma diversidade genética elevada em todas as populações analisadas neste estudo, inclusive nas populações ameríndias, que são, para a maioria dos locos estudados, pouco diversas. Sua diversidade se aproxima daquela que ocorre para o sistema HLA, uma das regiões mais polimórficas do genoma humano (PETZLER, LUZ e SOTOMAIOR, 1993; TSUNETO et al., 2003; OLIVEIRA, 2007; OLIVEIRA et al., 2008). Entretanto, quando comparadas às populações africana e eurobrasileira, essa diversidade é menor, o que pode ser explicado pela história evolutiva dessas populações.

A diversidade gênica, ou heterozigosidade esperada, baseada nas frequências alélicas observadas, é uma boa estimativa da diversidade genética nas populações humanas. Neste estudo, a heterozigosidade apresenta valor elevado em todas as populações, mostrando que este loco apresenta alto grau de polimorfismo.

Embora todas as populações difiram geneticamente entre si, a maior parte da diversidade genética é encontrada dentro das populações (entre indivíduos) e apenas pequena parte dela entre diferentes populações. 91,7% da diversidade do loco *KIR3DL1/S1* ocorre entre indivíduos e apenas 6,7% entre populações dentro dos grupos, isto é, entre os

Guarani Kaiowá, Nandeva e M'byá e entre os Kaingang de Ivaí e de Rio das Cobras. Os restantes 1,6% ocorrem entre os diferentes grupos ([AFRI], [EURO-BR], [ASI-BR], [Guarani] e [Kaingang]). Esses valores são muito próximos aos observados para o loco *MICA* (localizado na região de classe I do Complexo Principal de Histocompatibilidade, o *MHC*) (90,9 e 1%, respectivamente) avaliado apenas entre os ameríndios (excluindo africanos, eurobrasileiros e descendentes de asiáticos) (OLIVEIRA et al., 2008).

Vinte e cinco alelos foram identificados no conjunto total de indivíduos, incluindo quatro alelos ainda não descritos. As frequências destes alelos diferem significativamente entre todas as populações ($p < 10^{-5}$), até mesmo entre as populações ameríndias, que são historicamente mais próximas, como os Kaingang de Ivaí (KIV) e de Rio das Cobras (KRC) do estado do Paraná. Essas diferenças populacionais também foram observadas em outros estudos do nosso grupo, analisando microssatélites da região do *MHC* (OLIVEIRA, 2007) e os genes que codificam as moléculas coestimuladoras CD80 e CD86 (BELTRAME, 2008), e contrastam com estudos anteriores, que, analisando os locos *LMP2*, *TAP1*, *LMP7*, *TAP2* e outros genes do *MHC* de classe II e III, observaram grande homogeneidade entre os ameríndios e que consideraram KIV e KRC como um grupo único (PETZL-ERLER e McDEVITT, 1994; WEG-REMERS et al., 1997; FAUCZ, PROBST e PETZL-ERLER, 2000; TSUNETO et al., 2003).

A divergência bastante antiga entre os grupos humanos que deram origem às populações analisadas neste estudo pode explicar o elevado grau de diferenciação observado entre elas. Entretanto, para as populações ameríndias, que descendem de ancestrais comuns relativamente recentes, a diferenciação pode ser resultante de efeitos de fundador e da ação da deriva genética, uma vez que essas populações são pequenas e isoladas. Em vários estudos foi proposta a existência de fluxo gênico de não ameríndios para as populações ameríndias analisadas no presente trabalho (PETZL-ERLER, LUZ e SOTOMAIOR, 1993; FAUCZ, PROBST e PETZL-ERLER, 2000; TSUNETO et al., 2003; OLIVEIRA, 2007; OLIVEIRA et al., 2008). Dessa forma, o fluxo gênico recente com populações não indígenas, que pode ter ocorrido de forma diferenciada em cada população, também pode ter contribuído para essas diferenças.

Outros fatores, como a seleção natural, também podem ser responsáveis pelas diferenças genéticas observadas entre essas populações. Porém, entre os indígenas a ação de seleção natural de forma diferenciada é pouco provável, já que algumas dessas populações são geograficamente próximas e estiveram sujeitas, provavelmente, a condições ambientais semelhantes ao longo de sua história recente.

Com exceção dos novos alelos identificados, todos os demais são compartilhados entre duas ou mais populações deste estudo, ou com outras populações já descritas na literatura, incluindo populações de diferentes ancestralidades, evidenciando que esses alelos possuem uma origem evolutiva bastante antiga, antes da separação entre os principais grupos humanos. Os novos alelos identificados neste estudo ocorrem em grupos isolados, e têm, provavelmente, origem recente. O alelo *v.3DL1*309A* é comum às três populações Guaraní deste estudo (GKW, GND e GRC) e provavelmente surgiu antes da separação dessas tribos, ou então foi compartilhado por fluxo gênico entre esses grupos. Além disso, os novos alelos identificados nas populações indígenas deste estudo, provavelmente, são exclusivos destas populações, uma vez que não foram identificados nas várias populações europeias, asiáticas e africanas já investigadas. Apesar destes alelos não terem sido identificados em populações ameríndias da Venezuela (NORMAN et al., 2007), a investigação de outras populações pode esclarecer a origem e dispersão destes alelos dentro das Américas.

A menor diversidade genética nos ameríndios em relação as populações africana e eurobrasileira neste estudo pode ser observada pelo número de alelos nestas populações e, além disso, pelas frequências desses alelos, que diferem mais entre si nos ameríndios em comparação aos outros grupos estudados. Isso se deve provavelmente ao efeito fundador ou à deriva genética.

Apesar de diferirem geneticamente, os ameríndios e populações do leste asiático, apresentam uma maior similaridade, tanto no número de alelos identificados em cada população, quanto na distribuição das frequências desses alelos, quando comparadas aos outros dois grupos, o que se reflete em uma menor distância genética entre essas populações, como observado. Além disso, as populações ameríndias e asiáticas compartilham alelos com frequência relativamente alta que são inexistentes, ou pouco frequentes, em outros continentes, como o alelo *3DL1*02901*. Estas similaridades corroboram as hipóteses de que os ameríndios são originários de pequenos grupos vindos do leste asiático, há cerca de 20 mil anos (ROGERS et al., 1992; WALLACE e TORRONI, 1992; FAGUNDES, KANITZ e BONATTO, 2008; FAGUNDES et al., 2008).

O menor número de alelos foi observado nos indígenas Guaraní M'byá (GRC) e Kaingang de Rio das Cobras (KRC), nas quais foram encontrados seis alelos. O maior número foi verificado nos indígenas Guaraní Nandeva (GND), com 12 alelos. Apesar do número relativamente alto de alelos em algumas populações indígenas, nota-se que a maioria possui frequência extremamente reduzida. A frequência baixa, o não compartilhamento com outras populações indígenas, além da presença desses alelos em

européus e/ou africanos, nos quais esses alelos apresentam alta frequência, tornam muito provável a sua introdução por fluxo gênico recente de não indígenas. Entre os GND, população ameríndia com o maior número de alelos deste estudo, a taxa de miscigenação com não indígenas chega a 14% (TSUNETO et al., 2003). Entretanto, é possível que um ou outro desses alelos venha a ser identificado como pertencente ao *pool* gênico ancestral dos ameríndios e tenha estado presente nessas populações desde a sua fundação.

O alelo *3DL1*00401* provavelmente era inexistente nas populações ameríndias antes do contato com populações de outros continentes, ocorridos nos últimos cinco séculos, pois este alelo possui frequência relativamente alta na população eurobrasileira (10,8%) e africana (10,9%), mas foi identificado em um único indivíduo Guarani Ñandeva, não ocorrendo em nenhuma outra população indígena, nem nos descendentes de asiáticos (ASI-BR). A análise de maior número de populações ameríndias e asiáticas deverá esclarecer a origem e dispersão desse alelo.

Os valores de heterozigosidade observados nas populações ameríndias deste estudo são bastante similares aos observados para genes *HLA* de classe II (PETZL-ERLER e McDEVITT, 1994; TSUNETO et al., 2003) e para a maioria dos locos de microssatélites da região do *MHC* (classe I, II e III), avaliados por OLIVEIRA (2007). Por outro lado, a heterozigosidade do gene *KIR3DL1/S1* é superior a observada para os genes *TAP1*, *LMP2* e *LMP7*, também localizados na região do *MHC* de classe II (FAUCZ, PROBST e PETZL-ERLER, 2000).

A elevada heterozigosidade observada para o gene *KIR3DL1/S1*, os genes *HLA* e os microssatélites da região do *MHC*, mostra que essas regiões genômicas são duas das regiões mais polimórficas do genoma humano e sua diversidade é mantida nas populações humanas provavelmente por ação de seleção balanceadora. A elevada diversidade genética dessas regiões e a acentuada diferença entre as populações permite que sejam utilizadas como marcadores de ancestralidade em estudos populacionais.

Apesar de diferirem quanto à distribuição das frequências alélicas, foi observada uma pequena distância genética entre as duas populações Kaingang (KIV e KRC). Essa maior proximidade também é constatada quando se observam outros locos do genoma (PETZL-ERLER et al., 1993; PETZL-ERLER e McDEVITT, 1994; SALZANO et al., 1997; TSUNETO et al., 2003; OLIVEIRA et al., 2008). Uma pequena distância também é observada entre os grupos de Guarani (GKW, GND e GRC) e entre estes e as populações Kaingang. Essas distâncias reduzidas entre os grupos indígenas devem-se ao fato destas populações compartilharem uma origem evolutiva relativamente recente. Além disso, o fluxo

gênico entre grupos de ameríndios também pode explicar essa maior proximidade, uma vez que muitos desses grupos são culturalmente próximos.

A elevada frequência do alelo *KIR3DS1*01301* observada nas populações ameríndias deste estudo também foi observada em ameríndios da Venezuela, atingindo frequência de 51,6% nos Yucpa (NORMAN et al., 2007). Nos indígenas deste estudo, a maior frequência foi observada nos Guaraní M'byá (53%). Esse resultado, observado em todas as populações ameríndias já estudadas, revela que esse alelo foi trazido para o continente americano pelos paleoíndios. A elevada frequência do grupo de alelos ativadores nos ameríndios está de acordo com a sugestão de um aumento na frequência dos genes *KIR* ativadores nestas populações (GENDZEKHADZE et al., 2006; FLORES et al., 2007; RAJALINGAM et al., 2008; KULKARNI et al. 2008). O aumento da frequência de genes *KIR* ativadores em múltiplas populações indígenas possivelmente decorre de seleção natural sobre os diferentes genótipos *KIR*, já que os genes *KIR* ativadores participam da ativação das células *NK* que, por sua vez, desempenham um importante papel nas respostas imunológicas contra agentes infecciosos e células cancerígenas e na reprodução. Também de acordo com essas conclusões está o fato de o grupo alélico não funcional *004 não ocorrer em ameríndios, a não ser por fluxo gênico historicamente recente.

Durante as migrações humanas para fora da África, as populações sofreram efeito fundador e de deriva genética devido à redução do tamanho efetivo da população (efeito gargalo de garrafa), com consequente diminuição da sua diversidade genética (KIDD et al., 1998; LONJOU et al., 2003; CAVALLI-SFORZA e FELDMAN, 2003). Dessa forma, individualmente as populações apresentam, em geral, apenas uma parte da diversidade genética encontrada dentro do continente africano. Neste estudo, foram identificados 16 alelos em apenas 23 indivíduos africanos, o que demonstra a grande diversidade genética existente, de forma geral, nas populações africanas. Além disso, as frequências dos alelos identificados na população africana é relativamente baixa e similar para os diferentes alelos, mostrando que esses provavelmente coexistiram nesta população por longos períodos de tempo.

Os eurobrasileiros analisados neste estudo são uma população miscigenada, que, apesar da ancestralidade predominantemente européia, apresentam também contribuição ameríndia e africana. Estudos de nosso grupo, analisando genes *HLA* de classe I (*HLA-A*, -*B* e *C*) e II (*HLA-DR* e *-DQ*), nesse mesmo estrato populacional, mostraram que o componente europeu predomina (80,6%), sendo que a contribuição de africanos é de 12,5% e a de ameríndios é de 7,0%. Já para o estrato formado por indivíduos identificados como mulatos, o componente predominante foi o africano (49,5%), sendo as contribuições

européia de 41,8% e ameríndia de 8,7% (PROBST et al., 2000). BRAUN-PRADO e colaboradores (2000), analisando genes *HLA* de classe I, mostraram que não só para os “brancos”, mas também para os “mulatos”, o componente europeu é predominante. A contribuição de europeus, africanos e ameríndios foi, respectivamente, 94%, 3% e 3% para “brancos” e 57%, 39% e 4% para “mulatos”.

Essa origem heterogênea e a miscigenação podem explicar o grande número de alelos identificados nessa população (14 alelos) e suas frequências relativamente baixas. O alelo *3DS1*01301* é o mais frequente nessa população, que, de forma geral é o mais frequente também em outras populações já estudadas, excetuando as africanas e afrodescendentes (NORMAN et al., 2007). A contribuição africana para a constituição genética da população eurobrasileira pode ainda explicar a maior proximidade (menor distância genética) entre os africanos e eurodescendentes em relação às demais populações deste estudo.

Quando são observados os grupos alélicos relacionados às três linhagens ancestrais do gene *KIR3DL1/S1* (*005, *015 e *3DS1*013*), nota-se que na população africana há uma predominância dos alelos da linhagem *015, que somam a frequência de 78,2%. Esse dado sugere seleção positiva desta linhagem na população africana, como observado por Norman e colaboradores (2007) para outras populações africanas subsaarianas. A linhagem ativadora *3DS1*013* possui frequência bastante reduzida na população africana deste estudo, assim como naquelas analisadas por Norman e colaboradores (2007), e pode estar sendo mantida em baixa frequência por seleção, no continente africano.

Nos eurobrasileiros, as três linhagens ancestrais (*005, *015 e *3DS1*013*) possuem frequências relativamente próximas, sendo 30,4%, 33,3% e 24,5%, respectivamente, o que também se observa em outras populações européias e eurodescendentes já estudadas (<<http://www.allelefreqencies.net>>), indicando a manutenção destas linhagens por seleção balanceadora nestas populações. Já entre os ameríndios a linhagem *005 possui frequência bastante reduzida, assim como observado nos indígenas Yucpa e Bari da Venezuela (NORMAN et al., 2007). Essa frequência reduzida na maioria dos ameríndios estudados contrasta com a frequência de 26,3% observada nos indígenas Warao também na Venezuela (NORMAN et al., 2007), nos quais a frequência elevada pode ser explicada pelo efeito fundador, pela deriva genética ou pelo fluxo gênico com não ameríndios.

Além das três principais linhagens, outras linhagens recombinantes surgiram durante a história evolutiva deste gene. Um exemplo é a linhagem *001, que é comum à

maioria das populações e atinge a frequência de 12,1% nos indígenas Guaraní Nandeva e 10,8% nos eurobrasileiros. O compartilhamento entre diferentes populações e a frequência relativamente alta, permitem inferir que essa linhagem surgiu cedo durante a história evolutiva deste gene, mostrando que a evolução do gene *KIR3DL1/S1* é um processo constante, como também demonstrado pelo surgimento de novos alelos em populações autóctones.

A análise dos dendrogramas construídos a partir das distâncias genéticas entre as populações deste estudo e entre estas e outras populações já estudadas revela claramente as relações entre os principais grupos humanos, muito embora a história evolutiva de um único gene possa não representar, necessariamente, a história das populações humanas. As distâncias genéticas, calculadas a partir das frequências alélicas do gene *KIR3DL1/S1*, estão de acordo com a maior proximidade esperada entre as populações que são evolutivamente mais relacionadas. Os grandes grupos geográficos formam blocos bastante concisos, podendo ser definido um grupo ameríndio, um grupo do leste asiático, um grupo africano e um grupo europeu e eurodescendente.

6.2 DIVERSIDADE GENÉTICA DO LOCO *KIR3DL1/S1*: EVOLUÇÃO NEUTRA?

Os testes de neutralidade seletiva permitem verificar se há desvios para a hipótese de neutralidade seletiva sobre determinadas regiões genômicas, ou mesmo sobre todo genoma. A ação da seleção natural tem implicação direta sobre a diversidade nucleotídica de sequências genômicas, favorecendo a sua diversificação ou conservação.

A investigação exaustiva da sequência completa dos alelos encontrados nas populações deste estudo, bem como dos exons que codificam os diferentes domínios protéicos, separadamente, empregando diferentes testes, não permite rejeitar a hipótese de neutralidade seletiva sobre o gene *KIR3DL1/S1*. Entretanto, foram observados elevados valores (embora não significativos) para o teste D de Tajima em todas as populações. A população africana é uma exceção, pois apresenta valores negativos de D, o que está relacionado ao grande número de alelos nessa população, todos com frequência reduzida. O pequeno tamanho amostral da população africana também deve ser considerado na interpretação dos valores observados para os testes de neutralidade seletiva, pois pode conduzir a um viés amostral.

Esses valores elevados para o teste D de Tajima indicam a existência de vários alelos com frequências intermediárias nas populações deste estudo. Esses resultados

apontam para a mesma direção dos resultados obtidos por Norman e colaboradores (2007), que concluíram haver a ação de seleção balanceadora sobre a evolução deste gene na maioria das populações estudadas, excetuando-se a população africana, na qual há indícios de seleção direcional a favor da linhagem *015. A não significância para os valores do presente estudo deve-se, provavelmente ao reduzido tamanho amostral de cada população.

Quando avaliados apenas os alelos *3DL 1*, os valores encontrados para os testes de neutralidade seletiva foram bastante diferentes em comparação aos observados para o conjunto total de alelos, entretanto, continuam sendo não significativos.

A topologia do dendrograma construído a partir das frequências alélicas é outro indício de neutralidade seletiva, pois se assemelha a topologia de outros dendrogramas construídos a partir de marcadores neutros, aproximando as populações que são geograficamente mais próximas (Ramachandran et al., 2005). Além disso, a proporção dN/dS não difere nem quando analisada na sequência completa, nem dentro dos diferentes segmentos gênicos que correspondem a domínios estruturais e/ou funcionais da proteína, corroborando a hipótese de neutralidade seletiva sobre este gene. Por outro lado, a origem recente de mutações bem como a deriva genética, que ocorreu e possivelmente ainda ocorre principalmente nas populações ameríndias, podem ocultar as marcas provocadas pela seleção natural sobre esse loco.

A diversidade nucleotídica (π), tida como o número médio de diferenças nucleotídicas por sítio entre duas sequências e a variância amostral dessa diferença apresenta um valor elevado nas populações deste estudo, atingindo 13,4 na população eurobrasileira, mas não difere significativamente da média da diversidade observada para o genoma humano como um todo ($7,5 \times 10^{-4}$, Tishkoff e Varella, 2003). Também não há diferenças significativas para a comparação das populações deste estudo tomadas duas a duas. Esses valores elevados de π podem ser interpretados como ação de seleção balanceadora sobre a sequência do gene *KIR3DL1/S1*, desconsiderando a não significância observada. Essa falta de significância deve-se provavelmente ao tamanho das amostras analisadas. O menor valor de diversidade nucleotídica observado nos africanos deve-se ao fato do cálculo de π estar condicionado às frequências alélicas observadas e ser determinado principalmente por alelos (ou sítios) com frequência elevada, que é o que se observa nas populações com altos valores de π deste estudo. Nestas populações alguns alelos têm frequência bastante elevada enquanto que na população africana as frequências dos vários alelos são mais semelhantes entre si. Além disso, o tamanho da amostra populacional africana também tem implicância sobre o valor de π .

6.3 DIVERSIDADE FUNCIONAL DE RECEPTORES KIR3DL1/S1: SELEÇÃO BALANCEADORA?

Quando analisados apenas os resíduos de aminoácidos que interagem diretamente com as moléculas HLA, os receptores KIR formam motivos de ligação que agrupam vários alelos. Esses grupos estão relacionados às linhagens evolutivas deste gene.

Apesar de não termos observado desvios de neutralidade analisando a sequência completa dos alelos nas populações deste estudo, a análise dos códons que codificam os motivos de interação com HLA revela claramente a ação de seleção balanceadora e/ou positiva sobre essas regiões do gene *KIR3DL1/S1*. Nestas regiões há um acúmulo excessivo de substituições não sinônimas em relação às substituições sinônimas, em comparação com as demais regiões desse gene. A evolução dessas regiões funcionalmente relevantes sob ação de seleção balanceadora também pode ser verificada pelos valores positivos significativos para o teste de neutralidade seletiva D de Tajima, na região do exon 4 que contém esses códons, como observado nas populações ameríndias deste estudo.

Observando as frequências dos motivos de ligação, as diferenças populacionais são, de forma geral, menos acentuadas que as observadas pelas frequências alélicas. Esse fato também pode estar revelando a ação de seleção natural balanceadora, mantendo na população diferentes padrões estruturais/funcionais. Alelos que compartilham certo motivo funcional poderiam ser equivalentes. A topologia do dendrograma construído a partir das frequências dos motivos de ligação às moléculas HLA difere daquele construído pelas frequências alélicas, assim como observado por Norman e colaboradores (2007). Esta topologia não pode ser explicada por evolução neutra e é, portanto, mais um indício da ação de seleção natural sobre esses motivos funcionais.

Além disso, o cálculo do índice de fixação F_{st} sobre todos os sítios polimórficos encontrados ao longo das sequências dos alelos identificados nas populações deste estudo (48 sítios), revela que, de forma geral, os menores valores ocorrem para os sítios localizados nos exons que codificam os domínios extracelulares (D0, D1 e D2), que contém os resíduos de interação com HLA. Esse resultado demonstra que esses sítios são compartilhados entre os diferentes alelos nas populações e provavelmente são mantidos nas diferentes populações por seleção balanceadora, devido a sua importância funcional.

6.4 *KIR* E *HLA*: EVIDÊNCIAS DE COEVOLUÇÃO?

A importância da interação *KIR*:*HLA* na modulação da atividade das células *NK* sugere que estes genes podem ter evoluído conjuntamente (PARHAM, 2005). A acentuada diversificação observada nos motivos de interação dos receptores *KIR*, pelo acúmulo de substituições não sinônimas, com as moléculas *HLA*, que também são muito diversas, dá sustentação à hipótese de coevolução entre os dois locos. Essa diversificação nas regiões de interação pode resultar em combinações favoráveis entre *KIR* e *HLA*, de modo que a combinação dos genótipos desses dois locos pode ser importante no combate às infecções e na reprodução, sendo positivamente selecionadas (PANDO e PARHAM, 2005; NORMAN et al., 2007; MARTIN et al., 2007; THANANSHAI et al., 2007; CARR, THOMAS et al., 2008).

Outra evidência de que estes locos podem estar coevoluindo provém da análise de correlação de postos de Spearman, que revela uma correlação positiva entre *3DL1* e as moléculas *HLA* com epítipo Bw4, ou seja, a ação de seleção pode ter favorecido a coocorrência de receptor e ligante, mantendo interações funcionais entre *KIR* e *HLA*, que são extremamente importantes na regulação da atividade das células *NK*. A correlação negativa entre *3DS1* e Bw4 corrobora estudos recentes que demonstram que não há interação entre ligantes *HLA*-Bw4 e o receptor *KIR3DS1* (MORVAN et al., 2009).

Entretanto, avaliando a probabilidade de coocorrência dos diferentes genótipos *KIR3DL1/S1* com os epítipos Bw4 e Bw6, nenhum desvio significativo foi observado, aparentemente contrariando o resultado anterior, que indica uma relação entre *3DL1* e Bw4. Desse modo, conjuntamente esses dados sugerem que, apesar de haver coevolução por seleção, a ação da seleção não é suficientemente forte para manter as frequências de *3DL1* e *Bw4* estreitamente relacionadas. A correlação fraca entre *3DL1* e *Bw4*, observada através do teste de correlação de Pearson, também dá sustentação à hipótese de coevolução por seleção fraca.

6.5 DINÂMICA EVOLUTIVA DOS ALELOS *KIR3DL1/S1*

O polialelismo e polimorfismo do gene *KIR3DL1/S1* é um dos maiores observados para os genes *KIR*. A identificação de novos alelos e a coexistência destes com alelos já descritos em outras populações podem ser interpretados como resultado da ação de seleção positiva e balanceadora favorecendo a diversificação e manutenção desses alelos nas populações humanas. Entretanto, para o gene *KIR3DL1/S1* os resultados dos testes de

neutralidade seletiva nos levam a propor que a deriva (e o fluxo gênico) são os principais fatores evolutivos responsáveis pela diversidade alélica nas populações deste estudo.

A densidade média de SNPs em regiões exônicas para o genoma humano é de 1,14 SNP a cada 1 kb (The International HapMap Consortium, 2007). Considerando a sequência codificadora (1336 pb) do conjunto de alelos identificados nesse estudo, 48 polimorfismos foram identificados. Destes, 45 são polimorfismos de nucleotídeo único (SNPs). Somente na população africana foram 40 SNPs, um valor significativamente maior que o observado para o genoma como um todo ($p < 10^{-6}$). Essa densidade alta de SNPs também é observada nas populações ameríndias (número mínimo de SNPs = 30), sendo que os SNPs são em sua maioria compartilhados entre as populações. Esses resultados sugerem que há um acúmulo acelerado de substituições nucleotídicas durante a evolução deste gene, sendo que, provavelmente essas substituições ocorreram antes das migrações para fora do continente africano, uma vez que a maioria dos SNPs são compartilhados entre as populações.

Dos quatro novos alelos identificados neste estudo, três são resultado de mutações de ponto e um se deve provavelmente a um evento de conversão gênica entre um curto segmento no exon 5 de um alelo de uma das linhagens *3DL1* com outro da linhagem *3DS1*, originando um novo alelo *3DS1* (*v.3DS1*732C*) que compartilha um segmento polimórfico com alguns alelos *3DL1*. Essa grande contribuição de mutações pontuais na formação dos novos alelos identificados neste estudo contrasta com hipóteses anteriores, que afirmam que a recombinação teve um papel preponderante na diversificação dos gene *KIR3DL1/S1* (SCHILLING et al., 1998; GARDINER et al., 2001; NORMAN et al., 2009). Nossos dados sugerem uma ação igualmente importante das mutações de ponto para a evolução deste gene.

As variações dos novos alelos ocorrem no exon 3 e no exon 5, que codificam dois dos domínios de interação com HLA. Esses exons, juntamente com o exon 4, concentram o maior número de substituições nucleotídicas e codificam domínios com diferentes afinidades de interação com HLA. O exon 3 codifica o domínio zero (D0) dos receptores KIR e, embora esse domínio não participe diretamente da interação com as moléculas HLA, os polimorfismos presentes em D0 podem alterar a conformação dos domínios D1 e D2, interferindo na ligação com os ligantes. A transversoão (309C>A) (*v.3DL1*309A*), identificada no exon 3 nas populações Guaraní (GKW, GND e GRC) não ocasiona a substituição de aminoácidos na proteína e provavelmente essa variação não deve ter importância sobre a conformação dos domínios D1 e D2.

Das três transições identificadas para o exon 5, a transição G>A na posição 740 (códon 226) (*v.3DL1*740A*) leva a substituição do aminoácido arginina (Arg) por glutamina (Gln) dentro de um dos sítios de interação com as moléculas HLA no segundo domínio extracelular (D2) e foi identificada apenas na população africana. Essa substituição pode alterar o padrão de interação do receptor com seu ligante, uma vez que o aminoácido arginina é um aminoácido de caráter básico, enquanto a glutamina é um aminoácido neutro. Entretanto, para se entender qual a verdadeira implicação dessa substituição não sinônima no exon 5 é necessário o desenvolvimento de um estudo funcional que avalie a interação com os ligantes HLA. A identificação desse novo alelo em apenas um indivíduo deve-se, provavelmente, a sua origem recente dentro do continente africano, ou então, à sua ocorrência em grupos(s) étnico(s) ainda não analisado(s) diretamente ou através de seus descendentes em outros continentes que não a África.

O alelo *v.3DS1*857T* foi identificado em dois indivíduos da população Guarani Nandeva. Embora a transição 857C>T, no códon 265, exon 5, que ocasiona a substituição do aminoácido Treonina (Thr) por Isoleucina (Ile), não ocorra dentro do sítio de interação de KIR com HLA, essa substituição pode alterar a conformação do domínio 2 (D2), modificando o padrão de interação do receptor com o ligante. Aqui também a implicação funcional dessa substituição pode ser avaliada através de um estudo funcional.

A relação filogenética entre os alelos já descritos para este gene (considerando os novos alelos identificados nesse estudo) possui ramos que são sustentados por baixos valores de *bootstraps* o que se deve, provavelmente, a existência de muitos alelos intermediários entre os diferentes ramos. Nota-se, entretanto, que as três linhagens principais deste gene (**005*, **015* e *3DS1*013*) agrupam-se em três ramos diferentes, como demonstrado por Norman e colaboradores (2007), considerando apenas as sequências dos exons 3, 4 e 5. O maior valor de *bootstrap* que sustenta o ramo da linhagem *3DS1*013* (67) demonstra que esse grupo de alelos difere bastante dos alelos *3DL1*. A existência de poucos ramos dentro da linhagem *3DS1* demonstra a conservação que existe entre os alelos desta linhagem. A ação de seleção purificadora sobre *3DS1* pode explicar por que estes alelos são pouco variáveis, mesmo sendo evolutivamente tão antigos quanto *3DL1* (GARDINER et al., 2001; NORMAN et al., 2007). Além disso, pode-se observar através do dendrograma que existe uma grande distância entre os alelos humanos e a sequência homóloga nos chimpanzés, o que sustenta a hipótese de evolução acelerada e de forma espécie específica dos genes *KIR* (PARHAM et al., 2010; MOESTA et al., 2009; KELLEY, WALTER e TROWSDALE, 2005; KAKOO et al., 2000).

6.6 DESEQUILÍBRIO DE LIGAÇÃO

O desequilíbrio de ligação considerando os 43 SNPs ao longo das sequências dos alelos *3DL1* e *3DS1* (excluindo deleções) é o máximo possível para a maioria dos pares de SNPs, mesmo entre os mais distantes na sequência. Entretanto, os valores do coeficiente de correlação (r^2) são baixos na maioria dos casos, devido à frequências desproporcionais entre SNPs. O baixo valor de r^2 impossibilita a identificação de SNPs marcadores capazes de fornecer toda a informação do alelo em que ocorrem (*tag* SNPs). Um pequeno bloco de desequilíbrio com elevado valor de r^2 pode ser observado na extremidade 3' do gene *KIR3DL1/S1*, incluindo SNPs presentes nos exons 7, 8 e 9. Esses exons apresentam um grande número de diferenças entre as linhagens *3DL1* e *3DS1*, mas são praticamente idênticos dentro de cada linhagem, o que explica o elevado valor de r^2 nesta região.

Na população africana o desequilíbrio de ligação é menor que o observado nas demais populações, como historicamente se observa também para outros locos, uma vez que essas populações evoluem há muito mais tempo que as demais. Um menor desequilíbrio também é observado entre os eurobrasileiros e é resultado, provavelmente, da miscigenação desta população, além da maior antiguidade em relação aos ameríndios. Já entre os ameríndios, o desequilíbrio acentuado deve-se provavelmente a origem recente dessas populações e ao efeito fundador e à deriva genética, responsáveis pela baixa diversidade dessas populações. Esses padrões de desequilíbrio se assemelham aos observados em outros estudos do nosso grupo (OLIVEIRA, 2007; BELTRAME, 2008).

7 CONCLUSÕES

As populações investigadas neste estudo exibem uma elevada diversidade genética para o loco *KIR3DL1/S1*, sendo que esta diversidade é maior para os africanos, eurobrasileiros e asiáticos do que para as cinco populações indígenas estudadas. Essa diversidade provavelmente se deve a ação de seleção balanceadora sobre este gene nas diferentes populações. Além disso, as frequências alélicas diferem significativamente entre todas as populações, como resultado, da ação de diferentes fatores evolutivos ou efeitos demográficos sobre essas populações.

Os alelos pouco frequentes e não compartilhados entre as populações ameríndias, são resultado, provavelmente do fluxo gênico recente com não indígenas.

Apesar da grande diferenciação populacional, os indígenas são mais próximos entre si e mais distantes da população africana do que os europeus e os leste-asiáticos, o que está de acordo com a história evolutiva destas populações. Uma maior proximidade também é observada entre os ameríndios e os leste-asiáticos. A diversidade na população eurobrasileira se assemelha a observada em outras populações européias e eurodescendentes.

As populações indígenas possuem um menor grau de polimorfismo, com o alelo *3DS1*01301* tendo uma frequência bem elevada, como observado também nos indígenas da Venezuela (NORMAN et al., 2007). Esta observação corrobora estudos anteriores que sugerem um aumento na frequência dos genes *KIR* ativadores nas populações ameríndias, favorecendo sua sobrevivência, devido a importância destes genes na ativação das células *Natural Killer* (GENDZEKHADZE et al., 2006; FLORES et al., 2007; RAJALINGAM et al., 2008; KULKARNI et al. 2008). Já na população africana, o alelo *3DS1*01301*, possui frequência extremamente reduzida quando comparada as demais populações, sugerindo que, assim como proposto por Norman e colaboradores (2007), a linhagem *3DS1* tenha sido selecionada negativamente nesta população.

O desequilíbrio de ligação extremamente alto nas populações ameríndias é um indicativo da história evolutiva recente dessas populações, nas quais as frequências alélicas são resultantes principalmente da ação da deriva genética e do efeito fundador.

Foram identificados três novos alelos que resultam de mutações de ponto no exon 3 e no exon 5 do gene *KIR3DL1/3DS1* (*v.3DL1*309A*, *v.3DL1*740A* e *v.3DS1*857T*) e um novo alelo (*v.3DS1*732C*) na população GKW que, provavelmente, resulta de um evento

de conversão gênica entre um alelo do grupo 3DL1 e outro do grupo 3DS1. Esses novos alelos confirmam a grande diversidade dos genes *KIR* e corroboram as evidências de que esta é uma das regiões mais polimórficas do genoma humano. Além disso, a origem de três novos alelos por mutações pontuais sugere que este é um mecanismo evolutivo tão importante quanto a recombinação para este gene.

Aplicando diferentes testes evolutivos sobre a sequência completa do gene *KIR3DL1/S1* e também sobre os exons que codificam diferentes domínios do receptor, não foi possível descartar a hipótese de neutralidade seletiva neste gene. Essa evidência é sustentada pelo dendrograma populacional construído a partir das frequências alélicas, cuja topologia se assemelha a outros dendrogramas construídos a partir de marcadores neutros. Nossos resultados para os testes seletivos apontam na mesma direção dos resultados obtidos por Norman e colaboradores (2007) que sustentam evidências de seleção balanceadora sobre o gene *KIR3DL1/S1* na maioria das populações estudadas. A falta de significância para os testes aplicados neste estudo se deve ao reduzido tamanho amostral, o que dá pouco poder estatístico aos testes.

Avaliando apenas os códons nos exons quatro e cinco, que codificam os motivos de ligação às moléculas HLA, nós observamos um acúmulo significativo de substituições não sinônimas nos alelos encontrados nas populações do presente estudo. Nas populações ameríndias também foram obtidos valores positivos significativos para o teste D de Tajima nas regiões que incluem esses motivos no exon quatro. Além disso, a topologia do dendrograma construído a partir das frequências dos motivos de ligação à HLA nas populações deste estudo é diferente da obtida pelas frequências alélicas e não pode ser explicada por evolução neutra. Em conjunto, esses resultados mostram uma diversificação acentuada nas regiões que interagem diretamente com HLA devido, provavelmente, à ação de seleção natural e à coevolução dos locos *KIR* e *HLA*.

A correlação positiva observada entre *3DL1* e os motivos Bw4 nas populações deste estudo é uma evidência da ação de seleção mantendo interações funcionais KIR:HLA, que são importantes na modulação da atividade das células *NK*. Por outro lado, a correlação negativa entre *3DS1* e moléculas HLA com epítipo Bw4, sugere que estas moléculas não interagem.

8 REFERÊNCIAS BIBLIOGRÁFICAS

ABI-RACHED, L.; PARHAM, P. Natural selection drives recurrent formation of activating killer cell immunoglobulin-like receptor and Ly49 from inhibitory homologues. **The Journal of Experimental Medicine**, v. 201, p. 1319-1332, 2005.

ADAMS, E. J.; PARHAM, P. Species-specific evolution of MHC class I genes in the higher primates. **Immunological Review**, v. 183, p. 41-64, 2001.

ALTER, G.; MARTIN, M. P.; TEIGEN, N.; CARR, W. H.; SUSCOVICH, T. J.; SCHNEIDEWIND, A.; STREECK, H.; WARING, M.; MEIER, A.; BRANDER, C.; LIFSON, J. D.; ALLEN, T. M.; CARRINGTON, M.; ALTFELD, M. Differential natural killer cell – mediated inhibition of HIV-1 replication based on distinct KIR/HLA subtypes. **The journal of Experimental Medicine**, v. 204, p. 3027-3036, 2007.

ANDERSSON, S.; FAURIAT, C.; MALMBERG, J. A.; LJUNGGREN, H. G.; MALMBERG, K. J. KIR acquisition probabilities are independent of self-HLA class I ligands and increase with cellular KIR expression. **BLOOD**, v.114, p. 95-104, 2009.

ANFOSSI, N.; ANDRE, P.; GUIA, S.; FALK, C. S.; ROETYNCK, S.; STEWART, C. A.; BRESO, V.; FRASSATI, C.; REVIRON, D.; MIDDLETON, D.; ROMAGNE, F.; UGOLINI, S.; VIVIER, E. Human NK Cell Education by Inhibitory Receptors for MHC Class I. **Immunity**, v. 25, p. 331-342, 2006.

ASHIROVA, A. A.; MARTIN, M. P.; McVICAR, D. W.; CARRINGTON, M. The killer immunoglobulin-like receptor gene cluster: Tuning the genome defense. **Annual Reviews of Genomics and Human Genetics**, v. 7, p. 277-300, 2006.

BARRETT, J.C., FRY, B., MALLER, J.; DALY, M.J. Haploview: analysis and visualization of LD and haplotype maps. **Bioinformatics**, v. 21, p. 263, 2005.

BAXTER, G. A.; SMYTH, M. J. The role of NK cell in autoimmune disease. **Autoimmunity**, v. 35, p. 1-14, 2002.

BELICH, M. P.; MADRIGAL, J. A.; HILDEBRAND, W. H.; ZEMMOUR, J.; WILLIAMS, R. C.; LUZ, R.; PETZL-ERLER, M. L.; PARHAM, P. Unusual HLA-B alleles in two tribes of Brazilian Indians. **Nature**, v. 357, p. 326-329, 1992.

BELTRAME, M. H. **Estudo de populações de diferentes ancestralidades e evolução de polimorfismos dos genes CD80 e CD86**. Dissertação de Mestrado - Setor de Ciências Biológicas, Universidade Federal do Paraná, Curitiba, 2008.

BIASSONI, R.; CANTONI C.; MARRAS, D.; GIRON-MICHEL, J.; FALCO, M.; MORETTA, L.; DIMASI, N. Human Natural Killer cell receptors: insights into their molecular function e structure. **Journal of Cellular and Molecular Medicine**, v. 7, p. 376-387, 2003.

BONAGURA, V. R.; DU, Z.; ASHOURI, E.; LUO, L.; HATAM, L. J.; DEVOTI, J. A.; ROSENTHAL, D. W.; STEINBERG, B. M.; ABRAMSON, A. L.; GJERTSON, D. W.; REED, E. F.; RAJALINGAM, R. Activating killer cell immunoglobulin-like receptors 3DS1 and 2DS1 protect against developing the severe form of recurrent respiratory papillomatosis. **Human Immunology**, v. 71, p. 212-219, 2010.

BRAUN-PRADO, K.; MION, A. L. V.; PEREIRA, N. F.; CULPI, L.; PETZL-ERLER, M. L. HLA class I polymorphism, as characterised by PCR-SSOP, in a Brazilian exogamic population. **Tissue Antigens**, v.56, p. 417-427, 2000.

BROWN, M. G.; SCALZO, A. A.; YOKOYAMA, W. M. **The NKC and regulation of natural killer cell mediated immunity**. In MHC-Evolution Structure and Function, ed M. Kasahara; Tokyo:Springer, p. 287-304, 2000.

CALIGIURI, M. A. Human natural killer cells. **BLOOD**, v. 112, p. 461-469, 2008.

CARR, W. H.; PANDO, M. J.; PARHAM, P. KIR3DL1 Polymorphisms That Affect NK Cell Inhibition by HLA-Bw4 Ligand. **The Journal of Immunology**, v. 175, p. 5222-5229, 2005.

CAVALLI-SFOZA, L. L.; FELDMAN, M. W. The application of molecular genetic approaches to the study of human evolution. **Nature Genetics**, v. 33, p.266-275, 2003.

CHAN, H. W.; KURAGO, Z. B.; STEWART, C. A. ; WILSON, M. J.; MARTIN, M. P.; MACE, B. E.; CARRINGTON, M. TROWSDALE, J.; LUTZ, C. T. DNA Methylation Maintains Allele-specific *KIR* Gene Expression in Human Natural Killer Cells. **The Journal of Experimental Medicine**, v. 197, p. 2245-2255, 2003.

COOK, M. A.; MILLIGAN, D. W.; FEGAN, C. D.; DARBYSHIRE, P. J.; MAHENDRA, P.; CRADDOCK, C. F.; MOSS, P. A.; BRIGGS, D. C. The impact of donor KIR and patient HLA-C genotypes on outcome following HLA-identical sibling hematopoietic stem cell transplantation for myeloid leukemia. **BLOOD**, v. 15, p. 1521-1526, 2004.

COOLEY, S.; XIAO, F.; PITT, M.; GLEASON, M.; MCCULLAR, V.; BERGEMANN, T. L.; MCQUEEN, K. L.; GUETHLEIN, L. A.; PARHAM, P.; MILLER, J. S. A subpopulation of human peripheral blood NK cells that lacks inhibitory receptors for self-MHC is developmentally immature, **BLOOD**, v. 110, p. 578-586, 2007.

COOPER, M. A.; FEHNIGER, T. A.; CALIGIURI, M. A. The biology of human natural killer-cell subsets. **Trends in Immunology**, v. 22, p. 633-640, 2001.

CRUM, K. A.; LOGUE, S. E.; CURRAN, M. D.; MIDDLETON D. Development of a PCR-SSOP approach capable of defining the natural killer cell inhibitory receptor (KIR) gene sequence repertoires. **Tissue Antigens**, v. 56, p. 313-326, 2000.

DENIS, L.; SIVULA, J.; GOURRAUD, P. A.; KERDUDOU, N.; CHOUT, R.; RICARD, C; MOISAN, J. P.; GAGNE, K.; PARTANEN, J.; BIGNON, J. D. Genetic diversity of KIR natural killer cell markers in populations from France, Guadeloupe, Finland, Senegal and Reunion. **Tissue Antigens**, v. 66, p. 267-276, 2005.

Di SANTO, J. P. Natural Killer cell developmental pathways: A question of balance. **Annual Reviews Immunology**, v. 24, p. 257-286, 2006.

DRAGHI, M.; YAWATA, N.; GLEIMER, M.; YAWATA, M.; VALIANTE, N. M.; PARHAM, P. Single-cell analysis of the human NK cell response to missing self and its inhibition by HLA class I. **BLOOD**, v. 105, p. 2028-2035, 2005.

EWENS, W. J. The sampling theory of selectively neutral alleles. **Theoretical Population Biology**, v. 3, p. 87-112, 1972

EWING, B.; GREEN, P. Basecalling of automated sequencer traces using Phred II. Error probabilities. **Genome Research**, v.8, p.186-194, 1998.

EXCOFFIER, L.; LAVAL, G.; SCHNEIDER, S. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. **Evolutionary Bioinformatics Online**, v. 1, p. 47-50, 2005.

FAGUNDES, N. J.; KANITZ, R.; ECKERT, R.; VALLS, A. C.; BOGO, M. R.; SALZANO, F. M.; SMITH, D. G.; SILVA, W. A. Jr.; ZAGO, M. A.; RIBEIRO-DOS-SANTOS, A. K.; SANTOS, S. E.; PETZL-ERLER, M. L.; BONATTO, S. L. Mitochondrial population genomics supports a single pre-Clovis origin with a coastal route for the peopling of the Americas. **American Journal of Human Genetics**, v. 82, p. 583-592, 2008.

FAGUNDES, N. J. R.; KANITZ, R.; BONATTO, S. L. A reevaluation of the native American mtDNA genomic diversity and its bearing on the models of early colonization of Beringia. **PLOS One**, v. 3, 2008.

FAUCZ, F. R.; PROBST, C. M.; PETZL-ERLER, M. L. Polymorphism of LMP2, TAP1, LMP7 and TAP2 in Brazilian Amerindians and Caucasoids: implications for the evolution of allelic and haplotypic diversity. **European Journal of Immunogenetics**, v. 27, p. 5-16, 2000.

FAY, J. C.; WU, C. I. Hitchhiking under positive Darwinian selection. **Genetics**, v. 155, p. 1405-1413, 2000.

FELSENSTEIN, J. **PHYLIP (Phylogeny Inference Package) version 3.6**. Programa distribuido pelo autor. Department of Genome Sciences, University of Washington, Seattle, 2004.

FLORES, A. C.; MARCOS, C. Y.; PALADINO, N.; CAPUCCHIO, M.; THEILER, G.; ARRUVITO, L.; PARDO, R.; HABEGGER, A.; WILLIAMS, F.; MIDDLETON, D.; FAINBOIM, L. KIR genes polymorphism in Argentinean Caucasoid and Amerindian populations. **Tissue Antigens**, v. 69, p. 568-576, 2007.

FU, Y. X.; LI, W. H. Statistical tests of neutrality of mutations. **Genetics**, v. 133, p. 693-709, 1993.

FUNDAÇÃO NACIONAL DE SAÚDE (FUNASA). Disponível em: <
<http://www.funasa.gov.br/internet/index.asp>>. Acesso em 24/08/2010.

FUNDAÇÃO NACIONAL DO ÍNDIO (FUNAI). Disponível em: <<http://www.funai.gov.br>>. Acesso em: 24/08/2010.

GAO, X.; LIN, J.; WANG, L.; YU, L. Demethylating treatment suppresses natural killer cell cytolytic activity. **Molecular Immunology**, v. 46, p. 2064-2070, 2009.

GARDINER, C. M.; d Killer cell immunoglobulin-like receptors on NK cells: the how, where and why. **International Journal of Immunogenetics**, v. 35, p. 1-8, 2007.

GARDINER, C. M.; GUETHLEIN, L. A.; SHILLING, H. G.; PANDO, M.; CARR, W. H.; RAJALINGAM,; VILCHES, R. C. PARHAM, P. Different NK Cell Surface Phenotypes Defined by the DX9 Antibody Are Due to *KIR3DL1* Gene Polymorphism. **The Journal of Immunology**, v. 166, p. 2992-3001, 2001.

GATTAZZO, C.; MIORIN, M.; SCQUIZZATO, E.; TERAMO, A.; CABRELLE, A.; BALSAMO, M.; AGOSTINI, C.; VENDRAME, E.; FACCO, M.; ALBERGONI, M. P.; TRENTIN, L.; VITALE, M.; SEMENZATO, G.; ZAMBELLO, R. Lack of expression of inhibitory KIR3DL1 receptor in patients with NK-type lymphoproliferative disease of granular lymphocytes. **Haematologica**, v. 95, p. 1722-1729, 2010.

GAUDIERI, S.; DESANTIS, D.; MCKINNON, E.; MOORE, C.; NOLAN, D.; WITT, C. S.; MALLAL, S. A.; CHRISTIANSEN, F. T. Killer immunoglobulin-like receptors and HLA act both independently and synergistically to modify HIV disease progression. **Genes and Immunity**, v.6, p. 683-669, 2005.

GENDZEKHADZE, K.; NORMAN, P. J.; ABI-RACHED, L.; LAYRISSE, Z.; PARHAM, P. High KIR diversity in Amerindians is maintained using few gene-content haplotypes. **Immunogenetics**, v. 58, p. 474-480, 2006.

GENDZEKHADZE, K.; NORMAN, P. J.; ABI-RACHED, L.; GRAEF, T.; MOESTA, A. K.; LAYRISSE, Z.; PARHAM, P. Co-evolution of KIR2DL3 with HLA-C in a human population retaining minimal essential diversity of KIR and HLA class I ligands. **Proceedings of National Academy of Sciences**, v. 106, p. 18692-12697, 2009.

GEOBIS INTERNATIONAL. Disponível em: <http://www.geobis.com/mapas_africa_eng.html>. Acesso em: 20/07/2009.

GONZÁLEZ-JOSÉ, R.; BORTOLINI, M. C.; SANTOS, F. R.; BONATTO, S. L. The peopling of America: craniofacial shape variation on a continental scale and its interpretation from an interdisciplinary view. **American Journal of Physical Anthropology**, v. 137, p. 175-187, 2008.

GORDON, D.; ABAJIAN, C.; GREEN, P. **Consed: a graphical tool for sequence finishing.** **Genome Research**, v.8, p.195-202, 1998.

GREEN, P. **Phrap.** Disponível em: <www.genome.washington.edu/UWGC/analysisistools/phrap.htm>, 1994. Acesso em: 20/07/2009.

GRUMBERG, G.; MELIÀ, B. Guaraní Retã 2008: Povos Guaraní na Fronteira Argentina, Brasil e Paraguai. **CTI**, 2008.

GUETHLEIN, L. A.; FLODIN, L. R.; ADAMS, E. J.; PARHAM, P. NK Cell Receptors of the Orangutan (*Pongo pygmaeus*): A Pivotal Species for Tracking the Coevolution of Killer Cell Ig-Like Receptors with MHC-C. **The Journal of Immunology**, v. 169, p. 220-229, 2002.

GUETHLEIN, L. A.; OLDER AGUILAR, A. M.; ABI-RACHED, L.; PARHAM P. Evolution of Killer Cell Ig-Like Receptor (KIR) Genes: Definition of an Orangutan KIR Haplotype Reveals Expansion of Lineage III KIR Associated with the Emergence of MHC-C. **The Journal of Immunology**, v. 179, p. 491-504, 2007.

GUTIERREZ-RODRIGUEZ, M. E.; SANDOVAL-RAMIREZ, L.; DIAZ-FLORES, M.; MARSH, S. G.; VALLADARES-SALGADO, A.; MADRIGAL, J. A.; MEJIA-ARANGURE, J. M.; GARCIA, C. A.; HUERTA-ZEPEDA, A.; IBARRA-CORTES, B.; ORTEGA-CAMARILLO, C.; CRUZ, M. KIR Gene in ethnic and mestizo populations from Mexico. **Human Immunology**, v. 67, p. 85-93, 2006.

HALFPENNY, I. A.; MIDDLETON, D.; BARNETT, Y. A.; WILLIAMS, F. Investigation of Killer Cell Immunoglobulin-Like Receptor Gene Diversity: IV. KIR3DL1/S1. **Human Immunology**, v. 65, p. 602-612, 2004.

HALL, T. A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. **Nucleic Acids Symposium Series**, v. 41 p. 95-98, 1999.

HAO, L.; NEI, M. Rapid expansion of killer cell immunoglobulin-like receptor genes in primates and their coevolution with MHC class I genes. **Gene**, v. 347, p. 149-159, 2005.

HERSHBERGER, K. L.; KURIAN, J.; KORBER, B. T.; LETVIN, N. L. Killer cell immunoglobulin-like receptors (KIR) of the African-origin sabaeus monkey: evidence for recombination events in the evolution of KIR. **European Journal of Immunology**, v. 35, p. 922-935, 2005.

HERSHBERGER, K. L.; SHYAM, R.; MIURA, A.; LETVIN, N. L. Diversity of the killer cell Ig-like receptors of rhesus monkeys. **The Journal of Immunology**, v. 166, p. 4380-4390, 2001.

HIBY, S. E.; WALKER, J. J.; O'SHAUGHNESSY, K. M.; REDMAN, C. W. G.; CARRINGTON, M.; TROWSDALE, J.; MOFFETT, A. Combinations of maternal KIR and fetal HLA-C genes influence the risk of preeclampsia and reproductive success. **The Journal of Experimental Medicine**, v. 200, p. 957-965, 2004.

HOLLENBACH, J. A.; MEENAGH, A.; SLEATOR, C.; ALAEZ, C.; BENGOCHE, M.; CANOSSO, A.; CONTRERAS, G.; CREARY, L.; EVSEEVA, I.; GORODEZKY, C.; HARDIE, R. A.; HEMMING KARLSEN, T.; LIE, B.; LUO, M.; MARTINETTI, M.; NAVARETTE, C.; DE OLIVEIRA, D. C. M.; OZZELLA, G.; PASI, A.; PAVLOVA, E.; PINTO, S.; PORTO, L. C.; SANTOS, P.; SLAVCEV, A.; SRINAK, D.; TAVOULARI, S.; TONKS, S.; TRACHTENBERG, E.; VEJBAESYA, S.; MIDDLETON, D. Report from the killer immunoglobulin-like receptor (KIR) anthropology component of the 15th International Histocompatibility Workshop: worldwide variation in the KIR loci and further evidence for the co-evolution of KIR and HLA. **Tissue Antigens**, v. 76, p. 9-17, 2010.

HSU, K. C.; CHIDA, S.; GERAGHTY, D. E.; DUPONT, B. The killer cell immunoglobulin-like receptor (KIR) genomic region: gene order, haplotypes and allelic polymorphism. **Immunological Reviews**, v. 190, p. 40-52, 2002.

HUGHES, A. L. Evolution of the human killer cell inhibitory receptor family. **Molecular Phylogenetics and Evolution**, v. 25, p. 330-340, 2002.

HUGHES, A.L. **Adaptive Evolution of Genes and Genomes**. Oxford University Press, New York, 1999.

Instituto Brasileiro de Geografia e Estatística (IBGE). Disponível em: <<http://www.ibge.gov.br/home/default.php>>. Acesso em: 04/02/2011.

Instituto Socioambiental. Disponível em: <<http://www.socioambiental.org/website/pib/index.html>>. Acesso em: 24/08/2010.

IPD - KIR Database. Disponível em: <<http://www.ebi.ac.uk/ipd/kir/align.html>>. Acesso em: 24/08/2010.

JANEWAY, C. A.; TRAVERS, P.; WALPORT, M.; SHLOMCHIK, M. **Imunobiologia**. O sistema imune na saúde e na doença. 5. ed., Porto Alegre: Artmed, 2002.

JOBIM, M.; CHAGASTELLES, P.; SALIM, P. H.; PORTELA, P.; WILSON, T. J.; CURTI, A. G.; JOBIM, M. R.; JOÃO, D. A.; NARDI, N. B.; TSCHIEDEL, B.; JOBIM, L. F.; ROESLER, R.; SCHWARTSMANN, G. Association of killer cell immunoglobulin-like receptors and human leukocyte antigen-C genotypes in South Brazilian with type 1 diabetes. **Human Immunology**, v. 71, p. 799-803, 2010.

KARRE, K. NK cells, MHC class I molecules and the missing self. **Scandinavian Journal of Immunology**, v. 55, p. 221-228, 2002.

KELLEY, J.; TROWSDALE, J. Features of MHC and NK gene clusters. **Transplant Immunology**, v. 14, p. 129-134, 2005.

KELLEY, J.; WALTER, L.; TROWSDALE, J. Comparative genomics of natural killer cell receptor gene clusters. **PLOS Genetics**, v. 1, 2005.

KHAKOO, S. I.; GELLER, R.; SHIN, S.; JENKINS, J. A.; PARHAM, P. The D0 domain of KIR3D has a major histocompatibility complex class I binding enhancer. **The Journal of Experimental Medicine**, v. 196, p. 911-921, 2002.

KHAKOO, S. I.; THIO, C. L.; MARTIN, M. P.; BROOKS, C. R.; GAO, X.; ASTEMBORSKI, J.; CHENG, J.; GOEDERT, J. J.; VLAHOV, D.; HILGARTNER, M.; COX, S.; LITTLE, A. M.; ALEXANDER, G. J.; CRAMP, M. E.; O'BRIEN, S. J.; ROSENBERG, W. M. C.; THOMAS, D. L.; CARRINGTON, M. HLA and NK cell inhibitory receptor genes in resolving hepatitis C virus infection. **Science**, v. 305, p. 872-874, 2004.

KHAKOO, S. I.; RAJALINGAM, R.; SHUM, B. P.; WEIDENBACH, K.; FLODIN, L.; MUIR, D. G.; CANAVEZ, F.; COOPER, S. L.; VALIANTE, N. M.; LANIER, L. L.; PARHAM, P. Rapid Evolution of NK Cell Receptor Systems Demonstrated by Comparison of Chimpanzees and Humans. **Immunity**, v. 12, p. 687-698, 2000.

KIDD, K. K.; MORAR, B.; CASTIGLIONE, C. M.; ZHAO, H.; PAKSTIS, A. J.; SPEED, W. C.; BONNE-TAMIR, B.; LU, R. B.; GOLDMAN, D.; LEE, C.; NAM, Y. S.; GRANDY, D. K.; JENKINS, T.; KIDD, J. R. A global survey of haplotype frequencies and linkage disequilibrium at the DRD2 locus. **Human Genetics**, v. 103, p. 211-227, 1998.

KIM, S.; SUNWOO, J. B.; YANG, L.; CHOI, T.; SONG, Y. J.; FRENCH, A. R.; VLAHIOTIS, A.; PICCIRILLO, J. F.; CELLA, M.; COLONNA, M.; MOHANAKUMAR, T.; HSU, K. C.; DUPONT, B.; YOKOYAMA, W. M. HLA alleles determine differences in human natural killer cell responsiveness and potency. **Proceedings of the National Academy of Sciences of the United States of America**, v. 105, p. 3053-3058, 2007.

KOHLRAUSCH, F. B.; CALLEGARI-JACQUES, S. M.; TSUNETO, L. T.; PETZL-ERLER, M. L.; HILL, K.; HURTADO, A. M.; SALZANO, F. M.; HUTZ, M. H. Geography influences microsatellite polymorphism diversity in Amerindians. **American Journal of Physical Anthropology**, v. 126, n. 4, p. 463-470, 2005.

KULKARNI, S.; MARTIN, P.; CARRINGTON, M. The Yin and Yang of *HLA* and *KIR* in human disease. **Seminars in Immunology**, v. 20, p. 343-352, 2008.

KULKARNI, S.; SINGLE, M.; MARTIN, M. P.; RAJALINGAM, R.; BADWE, R. JOSHI, N.; CARRINGTON, M. Comparison of the rapidly evolving KIR locus in Parsis and natives of India. **Immunogenetics**, v. 60, p. 121-129, 2008.

LAHIRI, D. K.; NURNBERGER, J. I. Jr. A rapid non-enzymatic method for the preparation of DNA from blood for RFLP studies. **Nucleic Acids Research**, v. 19, p. 5444, 1991.

LENIER, L. L. NK cell recognition. **Annual Reviews Immunology**, v. 23, p. 225-274, 2005.

LI, H.; PASCAL, V.; MARTIN, M. P.; CARRINGTON, M.; ANDERSON, S. K. Genetic control of variegated *KIR* gene expression: polymorphisms of the bi-directional *KIR3DL1* promoter are associated with distinct frequencies of gene expression. **PLOS Genetics**, v. 4, 2008.

LJUNGGREN, H. G.; KARRE, K. In search of the 'missing self': MHC molecules and NK cell recognition. **Immunology Today**, v. 11, p. 237-244, 1990.

LONJOU, C.; ZHANG, W.; COLLINS, A.; TAPPER, W. J.; ELAHI, E.; MANIATIS, N.; MORTON, N. E. Linkage disequilibrium in human populations. **Proceedings of the National Academy of Sciences of the United States of America**, v. 100, p. 6069-6074, 2003.

LORENTZEN, A. R.; KARLSEN, T. H.; OLSSON, M.; SMESTAD, C.; MERO, I. L.; WOLDSETH, B.; SUN, J. Y.; SENITZER, D.; CELIUS, E. G.; THORSBY, E.; SPURKLAND, A.; LIE, B. A.; HARBO, H. F. Killer Immunoglobulin-like Receptor Ligand HLA-Bw4 Protects Against Multiple Sclerosis. **Annals Neurology**, v. 65, p. 658-666, 2009.

MARRERO, A. R.; SILVA-JUNIOR, W. A.; BRAVI, C. M.; HUTZ, M. H.; PETZL-ERLER, M. L.; RUIZ-LINARES, A.; SALZANO, F. M.; BORTOLINI, M. C. Demographic and evolutionary trajectories of the Guarani and Kaingang natives of Brazil. **American Journal of Physical Anthropology**, v. 132, p. 301-310, 2007.

MARTIN, A. M.; KULSKI, J. K.; GAUDIERI, S.; WITT, C. S.; FREITAS, E. M.; TROWSDALE, J.; CHRISTIANSEN, F. T. Comparative genomic analysis, diversity and evolution of two KIR haplotypes A and B. **Gene**, v. 335, p. 121-131, 2004.

MARTIN, M. P.; NELSON, G.; LEE, J. H.; PELLETT, F.; GAO, X.; WADE, J.; WILSON, M. J.; TROWSDALE, J.; GLADMAN, D.; CARRINGTON, M. Cutting edge: susceptibility to psoriatic arthritis: influence of activating killer Ig-like receptor genes in the absence of specific HLA-C alleles. **Journal of Immunology**, v. 169, p. 2818-2822, 2002.

MARTIN, M. P.; QI, Y.; GAO, X.; YAMADA, E.; MARTIN, J. N.; PEREYRA, F.; COLOMBO, S.; BROWN, E. E.; LESLEY, S. W.; PHAIR, J.; GOEDERT, J. J.; BUCHBINDER, S.; KIRK, D. G.; TELENTI, A.; CONNORS, M.; O'BRIEN, S. J.; WALKER, B. D.; PARHAM, P.; DEEKS, S. G.; MCVICAR, D. W.; CARRINGTON, M. Innate partnership of HLA-B and KIR3DL1 subtypes against HIV-1. **Nature Genetics**, v. 39, 2007.

MARTIN, M. P.; GAO, X.; LEE, J. H.; NELSON, G. W.; DETELS, R.; GOEDERT, J. J.; BUCHBINDER, S.; HOOTS, K.; VLAHOV, D.; TROWSDALE, J.; WILSON, M.; O'BRIEN, S. J.; CARRINGTON, M. Epistatic interaction between *KIR3DS1* and *HLA-B* delays the progression to AIDS. **Nature Genetics**, v. 31, p. 429-434, 2002.

MARTINEZ-BORRA, J.; KHAKOO, S. I. Speed and selection in the evolution of killer-cell immunoglobulin-like receptors. **International Journal Immunogenetics**, v. 35, p. 89-96, 2008.

<<http://www.allelefreqencies.net>>, MIDDLETON, D. 2003. Acesso em 25/01/2011.

MIDDLETON, D.; GONZELEZ, F. The extensive polymorphism of KIR genes. **Immunology**, v. 129, p. 8-19, 2009.

MILLER, M. P. RXC: a program for the analysis of contingency tables via the metropolis algorithm computer. **Department of Biological Sciences**, Northern Arizona University, Flagstaff, 1997.

MOESTA, A. K.; ABI-RACHED, L.; NORMAN, P. J.; PARHAM, P. Chimpanzees use more varied receptors and ligands than humans for Inhibitory KIR recognition of the MHC-C1 and MHC-C2 epitopes. **The Journal of Immunology**, v. 182, p. 3628-3637, 2009.

MORETTA, L.; MORETTA, A. Killer immunoglobulin-like receptors. **Current Opinion in Immunology**, v. 16, p. 626-633, 2004.

MORVAN, M.; WILLEM, C.; GAGNE, K.; Kerdudou, N.; DAVID, G.; SE'BILLE, V.; FOLLE'A, G.; BIGNON, J. D.; RETIE'RE, C. Phenotypic and Functional Analyses of KIR3DL1_ and KIR3DS1_ NK Cell Subsets Demonstrate Differential Regulation by Bw4 Molecules and Induced *KIR3DS1* Expression on Stimulated NK Cells. **The Journal of Immunology**, v. 182, p. 6727-6735, 2009.

National Center for Biotechnology Information (NCBI). Disponível em: <<http://www.ncbi.nlm.nih.gov/>> Acesso em: 24/08/2010.

NEI, M.; TAJIMA, F.; TATENO, Y. Accuracy of estimated phylogenetic trees from molecular data. **Journal of Molecular Evolution**, v. 19, p. 153-170, 1983.

NEVES, W. A.; HUBBE, M. Cranial morphology of early Americans from Lagoa Santa, Brazil: implications for the settlement of the New World. **Proceedings of the National Academy of Sciences of the United States of America**, v. 102, p. 18309-18314, 2005.

NEVES, W.A.; HUBBE, M.; PILÓ, L.B. Early Holocene human skeletal remains from Sumidouro Cave, Lagoa Santa, Brazil: history of discoveries, geological and chronological context, and comparative cranial morphology. **Journal of Human Evolution**, v. 52, p. 16-30, 2007.

NORMAN, P. J.; ABI-RACHED, L.; GENDZEKHADZE, K.; HAMMOND, J. A.; MOESTA, A. K.; SHARMA, D.; GRAEF, T.; MCQUEEN, K. L.; GUETHLEIN, L. A.; CARRINGTON, C. V. F.; CHANDANAYINGYONG, D.; CHANG, Y. H.; CRESPI, C.; SARUHAN-DIRESKENELI, G.; HAMEED, K.; KAMKAMIDZE, G.; KORAM, K. A.; LAYRISSE, Z.; MATAMOROS, N.; MILA, J.; PARK, M. H.; PITCHAPPAN, R. M.; DAN RAMDATH, D.; SHIAU, M. Y.; STEPHENS, H. A. F.; STRUIK, S.; TYAN, D.; VERITY, D. H.; VAUGHAN, R. W.; DAVIS, R. W.; FRASER, P. A.; RILEY, E. M.; RONAGHI, M.; PARHAM, P. Meiotic Recombination generates rich diversity in NK cell receptor genes, alleles, and haplotypes. **Genome Research**, v. 19, p. 757-769, 2009.

NORMAN, P. J.; ABI-RACHED, L.; GENDZEKHADZE, K.; KORBEL, D.; GLEIMER, M.; ROWLEY, D.; BRUNO, D.; CARRINGTON, C. V. F.; CHANDANAYINGYONG, D.; CHANG, Y-H.; CRESP, C.; SARUHAN-DIRESKENELI, G.; FRASER, P. A.; HAMEED, K.; KAMKAMIDZE, G.; KORAM, K. A.; LAYRISSE, Z.; MATAMOROS, N.; MILA, J.; PARK, M. H.; PITCHAPPAN, R. M.; DAN RAMDATH, D.; SHIAU, M-Y.; STEPHENS, H. A. F.; M.; STRUIK, S.; VERITY, D. H.; VAUGHAN, R. W.; TYAN, D.; DAVIS, R. W.; RILEY, E. M.; RONAGHI, M.; PARHAM, P. Unusual selection on the KIR3DL1/S1 natural killer cell receptor in Africans. **Nature Genetics**, v. 39, 2007.

NORMAN, P. J.; CARRINGTON, C. V.; BYNG, M.; MAXWELL, L. D.; CURRAN, M. D.; STEPHENS, H. A.; CHANDANAYINGYONG, D.; VERITY, D. H.; HAMEED, K.; RAMDATH, D. D.; VAUGHAN, R. W. Natural killer cell immunoglobulin-like receptor (*KIR*) locus profiles in African and South Asian populations. **Genes and Immunity**, v. 3, p. 86-95, 2002.

O'CONNOR, G. M.; GUINAN, K. J.; CUNNINGHAM, R. T.; MIDDLETON, D.; PARHAM, P.; GARDINER, C. M. Functional Polymorphism of the KIR3DL1/S1 Receptor on Human *NK* Cells. **The Journal of Immunology**, v. 178, p. 235-241, 2007.

Oligo Analyzer 3.0. Integrated DNA Technologies. Disponível em: <<http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/Default.aspx>>. Acesso em: 24/08/2010.

OLIVEIRA, L. **Diversidade Populacional de Microssatélites do Complexo Principal de Histocompatibilidade (MHC) em Populações Indígenas do Paraná e do Mato Grosso Do Sul.** Dissertação de Mestrado - Setor de Ciências Biológicas, Universidade Federal do Paraná, Curitiba, 2007.

OLIVEIRA, L. A.; RIBAS, F.; BICALHO, M. G.; TSUNETO, L. T.; PETZL-ERLER, M. L. High frequencies of alleles MICA*020 and MICA*027 in Amerindians and evidence of positive selection on exon 3. **Genes and Immunity**, v. 9, p. 697-705, 2008.

PAGE, R. D. M. TREEVIEW: An application to display phylogenetic trees on personal computers. **Computer Applications in the Biosciences**, v. 12, p.357-358, 1996.

PANDO, M. J.; GARDINER, C. M.; GLEIMER, M.; MCQUEEN, K. L.; PARHAM, P. The Protein Made from a Common Allele of *KIR3DL1*(*3DL1*004*) Is Poorly Expressed at Cell Surfaces due to Substitution at Positions 86 in Ig Domain 0 and 182 in Ig Domain. **The Journal of Immunology**, v. 171, p. 6640-6649, 2003.

PARHAM, P. The genetic and evolutionary balances in human NK cell receptor diversity. **Seminars in Immunology**, v. 20, p. 311-316, 2008.

PARHAM, P. Immunogenetics of killer cell immunoglobulin-like receptors. **Molecular Immunology**, v. 42, p. 459-462, 2005a.

PARHAM, P. MHC class I molecules and KIRs in human history, health and survival. **Nature Reviews Immunology**, v. 5, 2005b.

PARHAM, P. NK Cells and Trophoblasts: Patterns in Pregnancy. **The Journal of Experimental Medicine**, v. 200, p. 951-955, 2004.

PARHAM, P. Taking license with natural killer cell maturation and repertoire development. **Immunological Reviews**, v.214, p. 155-160, 2006.

PARHAM, P.; ABI-RACHED, L.; MATEVOSYAN, L.; MOESTA, A. K.; NORMAN, P. J.; OLDER AGUILAR, A. M.; Guethlein, L. A. Primate-specific regulation of natural killer cells. **Journal of Medical Primatology**, v.39, p.194-212, 2010.

PARHAM, P.; ARNETT, K. L.; ADAMS, E. J.; LITTLE, A. M.; TEES, K.; BARBER, L. D.; MARSH, S. G. E.; OHTA, T.; MARKOW, T.; PETZL-ERLER, M. L. Episodic evaluation and turnover of HLA-B in the indigenous human populations of the Americas. **Tissue Antigens**, v. 50, p. 219-232, 1997.

PARRA, F. C.; AMADO, R. C.; LAMBERTUCCI, J. R.; ROCHA, J.; ANTUNES, C. M.; PENA, S. D. J. Color and genomic ancestry in Brazilians. **Proceedings of the National Academy of Sciences of the United States of America**, v. 100, n. 1, p. 177-182, 2002.

PETZL-ERLER, M. L. Genetics of the immune responses and disease susceptibility. *Ciência e Cultura*, **Journal of the Brazilian Association for the Advancement of Science**, v.51, p.199-211, 1999.

PETZL-ERLER, M. L.; LUZ, R.; SOTOMAIOR, V. S. The HLA polymorphism of two distinctive South-American Indian tribes: The Kaingang and the Guarani. **Tissue Antigens**, v. 41, p. 227-237, 1993.

PETZL-ERLER, M. L.; MCDEVITT, H. O. Molecular analysis of the HLA-DRB genes in two tribes of Brazilian Indians. **Human Immunology**, v. 41, p.180-184, 1994.

PHYLIP. Felsenstein, J. Disponível em: <<http://evolution.genetics.washington.edu/phylip/getme.html>>. Acesso em: 20/07/2009.

PIMENTA, J. R.; ZUCCHERATO, L. W.; DEBES, A. A.; MASELLI, L.; SOARES, R. P.; MOURA-NETO, R. S.; ROCHA, J.; BYDŁOWSKI, S. P.; PENA, S. D. J. Color and Genomic Ancestry in Brazilians: A Study with Forensic Microsatellites. **Human Heredity**, v. 62, n. 4, p. 190-195, 2006.

Prefeitura Municipal de Curitiba (PMC). Disponível em: <<http://www.curitiba.pr.gov.br/>> Acesso em: 25/08/2010.

PROBST, C. M.; BOMPEIXE, E. P.; PEREIRA, N. F.; DE O DALALIO M. M.; VISENTAINER, J. E.; TSUNETO, L. T.; PETZL-ERLER, M. L. HLA polymorphisms and evaluation of European, African, and Amerindian contribution to the white and mulatto populations from Parana, Brazil. **Human Biology**, v. 72, p. 597-617, 2000.

PUCCIARELLI, H. M.; GONZÁLEZ-JOSÉ, R.; NEVES, W. A.; SARDI, M. L.; ROZZI, F. R. East-West cranial differentiation in pre-Columbian populations from Central and North America. **Journal of Human Evolution**, v. 54, p. 296-308, 2008.

RAJALINGAM, R.; HONG, M.; ADAMS, E. J.; SHUM, B. P.; GUETHLEIN, L. A.; PARHAM, P. Short KIR haplotypes in pygmy chimpanzee (*Bonobo*) resemble the conserved framework of diverse human KIR haplotypes. **The Journal of Experimental Medicine**, v. 193, 2001.

RAJALINGAM, R.; DU Z.; MEENAGH, A.; LUO, L.; KAVITHA, V. J.; PAVITHRA-ARULVANI, R.; VIDHYALAKSHMI, A.; SHARMA, S. K.; BALAZS, I.; REED, E. F.; PITCHAPPAN, R. M.; MIDDLETON, D. Distinct diversity of KIR genes in three southern Indian populations: comparison with world populations revealed a link between KIR gene content and pre-historic human migrations. **Immunogenetics**, v. 60, p. 207-217, 2008.

RAULET, D. H.; VANCE, R. E.; MCMAHON, C. W. Regulation of the natural killer cell receptor repertoire. **Annual Reviews Immunology**, v. 19, p. 291-330, 2001.

ROBERTSON, M. J. Role of chemokines in the biology of natural killer cells. **Journal of Leukocyte Biology**, v. 71, 2002.

RODRIGUEZ, R.; TYAN, D.; PARHAM, P. Allelic Polymorphism Synergizes with Variable Gene Content to Individualize Human *KIR* Genotype. **The Journal of Immunology**, v. 168, p. 2307-2315, 2002.

ROGERS, R. A.; ROGERS, L. A.; MARTIN, L. D. How the door opened: the peopling of the new world. **Human Biology**, v. 64, p. 281-302, 1992.

ROZAS, J.; SÁNCHEZ DELBARRIO, J.C.; MESSEGUER, X.; ROZAS, R. DnaSP, DNA polymorphism analyses by the coalescent and other methods. **Bioinformatics**, v.19 p. 2496-2497, 2003.

SAITOU, N.; NEI, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. **Molecular Biology and Evolution**, v. 4, p. 406–425, 1987.

SALTER, R. D.; CHEN, H.-W.; TADIKAMALLA, R.; LAWLOR, D. A. Domain organization and sequence relationship of killer cell inhibitory receptors. **Immunological Reviews**, v. 155, p. 175–182, 1997.

SALZANO, F. M.; CALLEGARI-JACQUES, S. M.; WEIMER, T. A.; FRANCO, M. H. L. P.; HUTZ, M. H.; PETZL-ERLER, M. L. Eletrophoretic protein polymorphisms in Kaingang and Guarani Indians of Southern Brazil. **American Journal of Human Biology**, v. 9, p. 505-512, 1997.

SAMBROOK, J. G.; BASHIROVA, A.; ANDERSEN, H.; PIATAK, M.; VERNIKOS, G. S.; COGGILL, P.; LIFSON, J. D.; CARRINGTON, M.; BECK, S. Identification of the ancestral killer immunoglobulin-like receptor gene in primates. **BMC Genomics**, 2006, v. 7, 2006.

SAMBROOK, J. G.; BASHIROVA, A.; PALMER, S.; SIMS, S.; TROWSDALE, J.; ABI-RACHED, L.; PARHAM, P.; CARRINGTON, M.; BECK, S. Single haplotype analysis demonstrates rapid evolution of the killer immunoglobulin-like receptor (*KIR*) loci in primates. **Genome Research**, v.15, p. 25-35, 2005.

SAMBROOK, J.; FRITSCH, E. F.; MANIATIS, T. **Molecular cloning - a laboratory manual**. 2. ed., Cold Spring Harbor laboratory Press, Plainview, 1989.

SANJANWALA, B.; DRAGHI, M.; NORMAN, P. J.; GUETHLEIN, L. A.; PARHAM, P. Polymorphic Sites Away from the Bw4 Epitope That Affect Interaction of Bw4_ HLA-B with KIR3DL1. **The Journal of Immunology**, v. 181, p. 6293-6300, 2008.

SCHILLING, H. G.; LIENERT-WIEDENBACH, K.; VALIANTE, N. M; UHRBERG, M.; PARHAM, P. Evidence for recombination as a mechanism for KIR diversification. **Immunogenetics**, v. 48, p. 413–416, 1998.

SINGLE, R. M.; MARTIN, M. P.; GAO, X.; MEYER, D.; YEAGER, M.; KIDD, J. R.; KIDD, K. K.; CARRINGTON, M. Global diversity and evidence for coevolution of KIR and HLA. **Nature Genetics**, 2007.

STEWART, C. A.; VAN BERGEN, J.; TROWSDALE, J. Different and divergent regulation of the KIR2DL4 and KIR3DL1 promoters. **The Journal of Immunology**, v. 170, p. 6073-6081, 2003.

SUZUKI, D. T.; GRIFFITHS, A. J. F.; MILLER, J. H.; LEWONTIN, R. C. **Introdução à Genética**. 6 ed., Rio de Janeiro, Guanabara Koogan, 1998.

TAJIMA, F. Statistical methods to test for nucleotide mutation hypothesis by DNA polymorphism. **Genetics**, v. 123, p. 585-595, 1989.

TAMURA, K.; DUDLEY, J.; NEI, M.; KUMAR, S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. **Molecular Biology and Evolution**, v.24, p.1596-1599, 2007.

TANIGUCHI, M.; KAWABATA, M. KIR3DL1/S1 genotypes and KIR2DS4 allelic variants in the AB KIR genotypes are associated with Plasmodium-positive individuals in malaria infection. **Immunogenetics**, v. 61, p. 717-730, 2009.

THANANCHAI, H.; GILLESPIE, G.; MARTIN, M. P.; BASHIROVA, A.; YAWATA, N.; YAWATA, M.; EASTERBROOK, P.; MCVICAR, D. W.; MAENAKA, K.; PARHAM, P.; CARRINGTON, M.; DONG, T.; ROWLAND-JONES, S. Cutting Edge: Allele-Specific and Peptide-Dependent Interactions between KIR3DL1 and HLA-A and HLA-B. **The Journal of Immunology**, v. 178, p. 33-37, 2007.

THE INTERNATIONAL HAPMAP CONSORTIUM. A second generation human haplotype map of over 3.1 million SNPs. **Nature**, v. 449, p. 851-862, 2007.

THOMAS, R.; YAMADA, E.; ALTER, G.; MARTIN, M. P.; BASHIROVA, A. A.; NORMAN, P. J.; ALTFELD, M.; PARHAM, P.; ANDERSON, S. K.; MCVICAR, D. W.; CARRINGTON, M. Novel *KIR3DL1* Alleles and Their Expression Levels on NK Cells: Convergent Evolution of *KIR3DL1* Phenotype Variation? **The Journal of Immunology**, v.180, p. 6743-6750, 2008.

THOMPSON, J. D.; HIGGINS, D. G.; GIBSON, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. **Nucleic Acids Research**, v.22, p. 4673-4680, 1994.

TISHKOFF, S. A.; REED, F. A.; FRIEDLAENDER, F. R.; EHRET, C.; RANCIARO, A.; FROMENT, A.; HIRBO, J. B.; AWOMOYI, A. A.; BODO, J. M.; DOUMBO, O.; IBRAHIM, M.; JUMA, A. T.; KOTZE, M. J.; LEMA, G.; MOORE, J. H.; MORTENSEN, H.; NYAMBO, T. B.; OMAR, S. A.; POWELL, K.; PRETORIUS, G. S.; SMITH, M. W.; THERA, M. A.; WAMBEBE, C.; WEBER, J. L.; WILLIAMS, S. M. The Genetic Structure and History of Africans and African Americans. **Science**, v. 324, p. 1035-1044, 2009.

TISHKOFF, S. A.; VERRELLI, B. C. Patterns of Human Genetic Diversity: Implications for Human Evolutionary History and Disease. **Annual Reviews Genomics of Human Genetics**, v. 4, p. 293-340, 2003.

TONEVA M.; LLEAPFAAYGE, V. G.; DULPHY, N.; BUSSON, M.; LESTER, S.; VU-TRIEN, A.; MICHAYLOVA, A.; NAUMOVA, E.; MCCLUSKEY, J.; CHARRON, D. Brief communication: Genomic diversity of natural killer cell receptor genes in three populations. **Tissue Antigens**, v. 57, p. 358-362, 2001.

TRAHERNE, J. A.; MARTIN, M.; WARD, R.; OHASHI, M.; PELLETT, F.; GLADMAN, D.; MIDDLETON, D.; CARRINGTON, M.; TROWSDALE, J. Mechanisms of copy number variation and hybrid gene formation in the KIR immune gene complex. **Human Molecular Genetics**, v. 19, p. 737-751, 2010.

TROWSDALE, J. Genetic and functional relationship between MHC and NK receptor genes. **Immunity**, v. 15, p. 363-374, 2001.

TROWSDALE, J.; MOFFETT, A. NK receptor interactions with MHC class I molecules in pregnancy. **Seminars in Immunology**, v. 20, p. 317-320, 2008.

TSUNETO, L. T.; PROBST, C. M.; HUTZ, M. H.; SALZANO, F. M.; RODRIGUEZ-DELFIN, L. A.; ZAGO, M. A.; HILL, K.; HURTADO, A. M.; RIBEIRO-DOS SANTOS, A. K. C.; PETZL-ERLER, M. L. HLA class II diversity in seven Amerindian populations. Clues about the origins of the Aché. **Tissue Antigens**, v. 62, p. 512-526, 2003.

UHRBERG, M.; VALIANTE, N. M.; YOUNG, N. T.; LANIER, L. L.; PHILLIPS, J. H. PARHAM, P. The repertoire of killer cell Ig-like receptor and CD94:NKG2A receptors in T cells: clones sharing identical $\alpha\beta$ TCR rearrangement express highly diverse killer cell Ig-like receptor patterns. **The Journal of Immunology**, v. 166, p. 3923-3932, 2001.

van der SLIK, A. R.; KOELMAN, B. P.; VERDUIJIN, W.; BRUINING, G. J.; ROEP, B. O.; GIPHART, M. J. KIR in type I diabetes: disparate distribution of activating and inhibitory natural killer cells receptor in patients *versus* HLA-matched control subjects. **Diabetes**, v. 52, p. 2639-2642, 2003.

VARLA-LEFTHERIOTI, M.; SPYROPOULOU-VLACHOU, M.; NIKOU, D.; KERAMITSOGLOU, T.; DARLAMITSOU, A.; TSEKOURA, C.; PAPADIMITROPOULOS, M.; LEPAGE, V.; BALAFOUTAS, C.; STAVROPOULOS- GOKAS, C. Natural killer (NK) cell receptors' repertoire in couples with recurrent spontaneous abortions. **American Journal of Reproductive Immunology**, v. 49, p.183-191, 2003.

VELARDI, A. Role of KIRs and KIR ligands in hematopoietic transplantation. **Current Opinion in Immunology**, v. 20, p. 581-587, 2008.

VILCHES, C.; PARHAM, P. KIR: Diverse, rapidly evolving receptors of innate and adaptive immunity. **Annual Reviews Immunology**, v. 20, p. 217-251, 2002.

VIVIER, E.; NUNÈS, J. A.; VÉLY, F. Natural Killer cell signaling pathways. **Science**, v. 306, 2004.

WALLACE, D. C.; TORRONI, A. American Indian prehistory as written in the mitochondrial DNA: a review. **Human Biology**, v. 64, p. 403-416, 1992.

WAN, A. M.; ENNIS, P.; PARHAM, P.; HOLMES, N. The primary structure of HLA-A32 suggests a region involved in formation of the Bw4/Bw6 epitopes. **The Journal of Immunology**, v. 137, p. 3671-3674, 1986.

WANG, S.; LEWIS Jr, C. M.; JAKOBSSON, M.; RAMACHANDRAN, S.; RAY, N.; BEDOYA, G.; ROJAS, W.; PARRA, M. V.; MOLINA, J. A.; GALLO, C.; MAZZOTTI, G.; POLETTI, G.; HILL, K.; HURTADO, A. M.; LABUDA, D.; KLITZ, W.; BARRANTES, R.; BORTOLINI, M. C.; SALZANO, F. M.; PETZL-ERLER, M. L.; TSUNETO, L. T.; LLOP, E.; ROTHHAMMER, F.; EXCOFFIER, L.; FELDMAN, M. W.; ROSENBERG, N. A.; RUIZ-LINARES, A. Genetic Variation and Population Structure in Native Americans. **PloS Genetics**, v. 3, p. 2049-2067, 2007.

WATTERSON, G. The homozygosity test of neutrality. **Genetics**, v. 88, p. 405-417, 1978.

WHANG, D. H.; PARK, H.; YOON, J. A.; PARK, M. H. Haplotype Analysis of Killer Cell Immunoglobulin-Like Receptor Genes in 77 Korean Families. **Human Immunology**, v. 66, p. 146-154, 2005.

WEG-REMER, S.; BRENDEN, M.; SCHWARZ, E.; WITZEL, K.; SCHNEIDER, P. M.; GUERRA, L. K.; REHFELDT, I. R.; LIMA, M. T.; HARTMANN, D.; PETZL-ERLER, M. L.; de MESSIAS, I. J.; MAUFF, G. Major histocompatibility complex (MHC) class III genetics in two Amerindian tribes from Southern Brazil: the Kaingang and the Guarani. **Human Genetics**, v. 100, p. 548-556, 1997.

WILSON, M. J.; TORKAR, M.; HAUDE, A.; MILNE, S.; JONES, T.; SHEER, D.; BECK, S.; TROWSDALE, J. Plasticity in the organization and sequences of human KIRyILT gene families. **Proceedings of National Academy of Sciences**, v. 97, p. 4778-4783, 2000.

YAWATA, M. N.; YAWATA, N.; DRAGHI, M.; LITTLE, A-M.; PARTHENIOU, F.; PARHAM, P. Roles for HLA and KIR polymorphisms in natural killer cell repertoire selection and modulation of effector function. **The Journal of Experimental Medicine**, v. 203, p. 633-645, 2006.

YAWATA, M.; YAWATA, N.; DRAGHI, M.; PARTHENIOU, F.; LITTLE, A. M.; PARHAM, P. MHC class I-specific inhibitory receptors and their ligands structure diverse human NK-cell repertoires toward a balance of missing self-response. **Blood**, v. 112, p. 2369-2380, 2008.

YEN, J. H.; MOORE, B. E.; NAKAJIMA, T.; SCHOLL, D.; SCHAID, D. J.; WEYAND, C. M.; GORONZY, J. J. Major Histocompatibility Complex class I recognizing receptors are disease risk genes in rheumatoid arthritis. **Journal of Experimental Medicine**, v. 193, p. 1159-1168, 2001.

YOKOYAMA, W. M.; KIM, S. Licensing of natural killer cells by self-major histocompatibility complex class I. **Immunological Reviews**, v. 215, p. 143-154, 2006.

YU, J.; HELLER, G.; CHEWNING, J.; KIM, S.; YOKOYAMA, W. M.; HSU, K. C. Hierarchy of the Human Natural Killer Cell Response Is Determined by Class and Quantity of Inhibitory Receptors for Self-HLA-B and HLA-C Ligands. **The Journal of Immunology**, v. 179, p. 5977-5989, 2007.

YUSA, S.; CATINA, T. L.; CAMPBELL, K. S. SHP-1 and Phosphotyrosine-Independent Inhibitory Signaling by a Killer Cell Ig-Like Receptor Cytoplasmic Domain in Human NK Cells. **The Journal of Immunology**, v. 168, p. 5047-5057, 2002.

9 APÊNDICE: MATERIAIS E MÉTODOS

9.1 CARACTERIZAÇÃO DAS AMOSTRAS

Neste estudo foram analisados 451 indivíduos, pertencentes a oito populações distintas. As frequências alélicas do gene *KIR3DL1/S1* foram determinadas para uma amostra da população africana, de indivíduos originários dos países de Angola, Gana, Guiné-Bissau e Moçambique (AFRI, n=23), para a população urbana de Curitiba e região metropolitana (EURO-BR, n=51), para uma amostra de brasileiros descendentes de asiáticos (ASI-BR, n=56) e para cinco populações ameríndias pertencentes às tribos indígenas Guarani M'byá (GRC, n=67), da reserva indígena Rio das Cobras, município de Nova Laranjeiras - PR, Guarani Kaiowá (GKW, n=64), dos municípios de Limão Verde e Amambaí, Mato Grosso do Sul, Guarani Ñandeva (GND, n=58), dos municípios de Amambaí e Porto Lindo, MS, e Kaingang, das aldeias indígenas Rio das Cobras (KRC, n=72) e Ivaí (KIV, n=60), município de Manoel Ribas, PR (FIGURA 32). As amostras foram coletadas a partir da década de 1980 pela equipe do LGMH.

Alguns indivíduos das populações EURO-BR, GRC, GKW, GND e KRC foram selecionados de acordo com o haplótipo de genes KIR, cujas tipagens foram realizadas anteriormente por nosso grupo (LGMH - Laboratório de Genética Molecular Humana). Foram escolhidos os indivíduos portadores de perfis únicos dos genes *KIR*, ou seja, portadores de haplótipos identificados uma única vez, e a amostra foi completada aleatoriamente.



FIGURA 32. Mapa da América do Sul, indicando a localização das populações analisadas neste estudo. 1 - área de reserva Ivaí, município de Manoel Ribas, PR; 2 - área de reserva Rio das Cobras, município de Nova Laranjeiras, PR; 3 - áreas Amambaí e Limão Verde, município de Amambaí, MS; 4: área de reserva indígena Porto Lindo, MS; 5 - região metropolitana de Curitiba. **Fonte:** OLIVEIRA et al., 2008.

A amostra da população oriental é composta principalmente de indivíduos descendentes de japoneses vindos para o Brasil nos últimos 100 anos. Já a amostra populacional africana é formada por indivíduos nascidos em quatro países da África (FIGURA 33), em sua maioria estudantes que vieram concluir seus estudos na cidade de Curitiba. Alguns indivíduos, originários de Moçambique, foram obtidos em colaboração com o Dr. Guilherme Suarez-Kurtz da Divisão de Farmacologia, INCA, (Rio de Janeiro). A amostra da população brasileira não indígena é representada pela população residente em Curitiba e região metropolitana, e foi coletada a partir de 1994 pela equipe do LGMH.

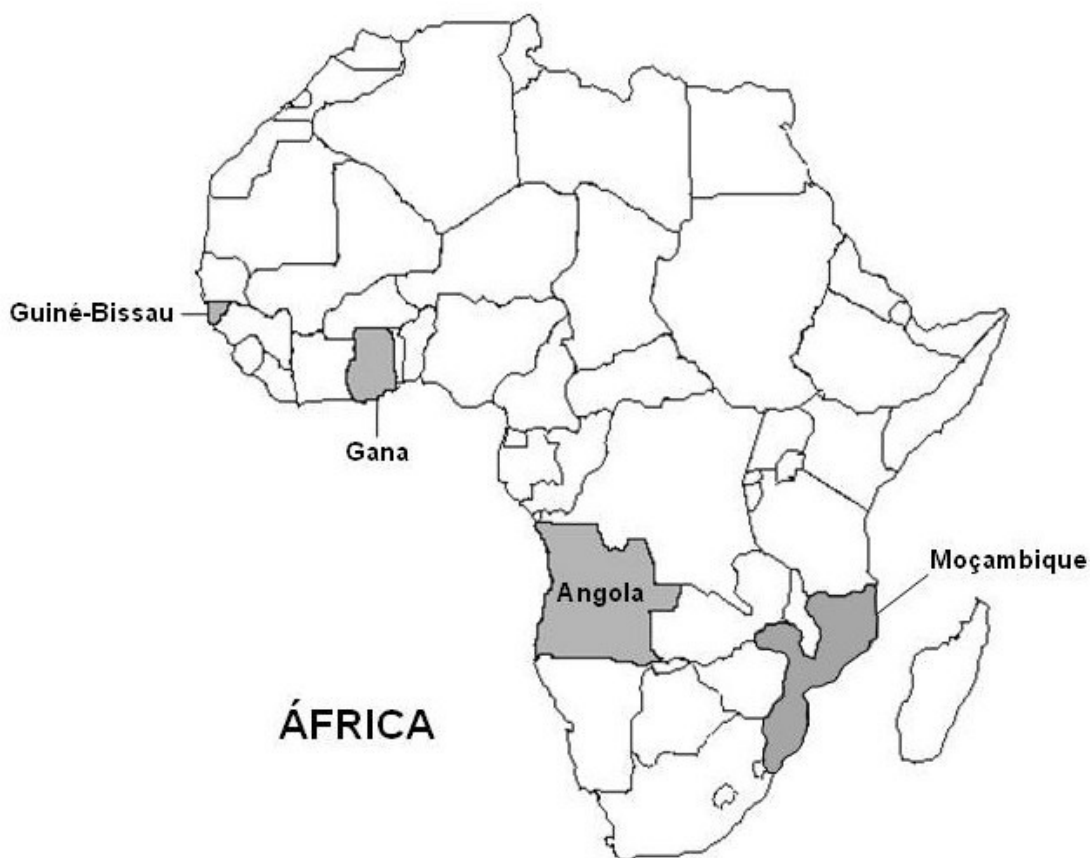


FIGURA 33. Mapa da África. Em destaque os países de origem dos africanos analisados neste trabalho: Moçambique, Guiné-Bissau, Gana e Angola. **Fonte:** adaptado de <<http://clio.rediris.es/actividades/africa/africa.htm>>.

O DNA foi extraído a partir do sangue total, através dos métodos fenol-clorofórmio-álcool isoamílico (SAMBROOK et al., 1989, com protocolo adaptado do 13º IHWC (International Histocompatibility Workshop and Congress, 2000) ou salting-out (LAHIRI e NURNBERGER, 1991). Os indivíduos foram voluntários e assinaram um termo de consentimento para participação em pesquisa. Este trabalho foi aprovado pelo Comitê de Ética em Pesquisa com Seres Humanos do SCB (Setor de Ciências Biológicas) da Universidade Federal do Paraná e pela CONEP (Comissão Nacional de Ética em Pesquisa). A investigação dos indígenas foi autorizada pela Fundação Nacional do Índio (FUNAI) e pelos líderes das tribos.

9.2 AMPLIFICAÇÃO E SEQUENCIAMENTO

As regiões correspondentes aos exons 1-5, sete e nove do gene *KIR3DL1/S1* foram amplificadas por *PCR* (Reação em Cadeia da Polimerase) a partir do DNA genômico. Após a amplificação, os produtos foram submetidos a técnica de sequenciamento. Os exons três, quatro e cinco foram sequenciados em ambas as direções, enquanto que os exons um, dois, sete e nove apenas em uma única direção (5'-3').

Os oligonucleotídeos iniciadores para a reação de *PCR* e também para as reações de sequenciamento foram escolhidos com base na sequência do gene disponível no banco de dados IPD-KIR Database (<http://www.ebi.ac.uk/ipd/kir/align.html>), através da análise das sequências escolhidas no programa Oligo Analyzer[®] 3.0 e são mostrados na tabela 17, onde também são mostrados os tamanhos dos fragmentos a serem amplificados. Os oligonucleotídeos para amplificação dos exons 3 e 4 foram adaptados do trabalho de Halfpenny e colaboradores (2004) e os utilizados para os exons 1 e 2, obtidos em colaboração com o National Cancer Institute, Frederick - USA, onde também foi realizado parte do sequenciamento dos exons um, dois e nove. Os demais oligonucleotídeos foram projetados a partir das sequências dos genes *KIR3DL1* (**NG_005990.1**) e *KIR3DS1* (**NG_005994.1**), disponíveis no Genbank (NCBI – National Center for Biotechnology Information). Os oligonucleotídeos iniciadores utilizados para o sequenciamento do exon quatro foram os mesmos que os utilizados para a amplificação. Para o sequenciamento do exon cinco, o oligonucleotídeo 3' foi o mesmo que o utilizado na amplificação.

TABELA 17. Oligonucleotídeos iniciadores para as reações de amplificação e sequenciamento do gene *KIR3DL1/S1*.

| | Localização | Sequência (5' - 3') | TM |
|---|------------------|--|---------|
| Exons 1e2 - Peptídeo Sinal | | | |
| Amplificação | 5' UTR | KIR3DL1/S1_EX1e2FOR_A: CCTGGGAGTTTAAATCATTTG | 49,4 °C |
| | exon 3 | KIR3DL1/S1_EX1e2REV_A: CTGTGACCATGATCACCAC | 53,1 °C |
| Sequenciamento | | KIR3DL1/S1_EX1FOR_S: TTTAGGCATCTCGTGTTCGG | |
| | | KIR3DL1/S1_EX2FOR_S: AGTGGGGGCAGCAGGGTG | |
| Exon 3 - D0 | | | |
| Amplificação | exon 2 | KIR3DL1/S1EX3FOR_A: TTGGTCCAGAGGCCGGTC | 62,3°C |
| | exon 3 | KIR3DL1/S1EX3REV_A: CCTGTGACCATGATCACCAY | 55,1 °C |
| Sequenciamento | | KIR3DL1/S1EX3FOR_S: CCATGCCGTGTCTACTTTGT | |
| | | KIR3DL1/S1EX3REV_S: CTGTGACCATGATCACCA | |
| Exon 4 - D1 | | | |
| Amplificação | exon 4, intron 3 | KIR3DL1/S1EX4FOR_A: TCCCCTGGTGAAATCAGGAGA | 57,7 °C |
| | exon 4 | KIR3DL1/S1EX4REV_A: GTGACCACGATGTCCAGGGGA | 61,2 °C |
| Sequenciamento | | KIR3DL1/S1EX4FOR_S: TCCCCTGGTGAAATCAGGAGA | |
| | | KIR3DL1/S1EX4REV_S: GTGACCACGATGTCCAGGGGA | |
| Exon 5 - D2 | | | |
| Amplificação | exon 4 | KIR3DL1/S1EX5FOR_A: GCTACGGTTCTGTTACTCACA | 54,4 °C |
| | exon 5 | KIR3DL1/S1EX5REV_A: TGACCAYTCGTAGGGAGAGTG | 56,9 °C |
| Sequenciamento | | KIR3DL1/S1EX5FOR_S: GGGAGTGAGTTCTCAGCTCA | |
| | | KIR3DL1/S1EX5REV_A: TGACCAYTCGTAGGGAGAGTG | |
| Exon 7e9 - Domínio Transmembrana e Citoplasmático | | | |
| Amplificação | exon 7 | KIR3DL1_Ex9FOR_A: AGTGGTCATCATCCTTTCATC | 54,0 °C |
| | exon 9 | KIR3DL1_Ex9REV_A: GTGTACAAGATGGTATCTGTA | 48,9 °C |
| Sequenciamento | | KIR3DL1/S1_Ex7_S: GTGCTTGTCKAAAGAGAYGC | |
| | | KIR3DL1/S1_Ex9_S: CTGCAGGGAACAGAACAGCC | |

Y - C/T; K - G/T; TM - Temperatura de fusão calculada pelo programa Oligo Analyser (versão 3.1. Disponível em: <<http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/>>).

Os protocolos para as reações de amplificação foram:

1) Exons 1 e 2 - tampão sulfato 1 X, 1,5 mM de cloreto de magnésio ($MgCl_2$), 0,2 mM de dNTP, 0,4 M de cada oligonucleotídeo iniciador, 1 U de Taq DNA polimerase (*Taq Platinum*, Invitrogen®) e 40 ng de DNA genômico, em um volume final de 20 μ l. As condições de ciclagem foram: a - desnaturação inicial a 94 °C por 3 minutos; b - 40 ciclos de desnaturação a 94 °C por 30 seg, hibridação a 58,5 °C por 30 seg e extensão a 72 °C por 2 min e 30 seg e; c - uma etapa final de alongamento, a 72 °C por 10 minutos.

2) Exon 3 - tampão sulfato 1 X, 2,5 mM de cloreto de magnésio ($MgCl_2$), 0,2 mM de dNTP, 0,2 M de cada oligonucleotídeo iniciador, 0,5 U de Taq DNA polimerase (*Taq Platinum*, Invitrogen®) e 20 ng de DNA genômico, em um volume final de 10 μ l. As condições de ciclagem foram: a - desnaturação inicial a 96 °C por 5 minutos; b - 15 ciclos de desnaturação a 96 °C por 1 min, hibridação a 65 °C por 45 seg e extensão a 72 °C por 1 min; c - 30 ciclos de desnaturação a 96 °C por 1 min, hibridação a 62 °C por 45 seg e extensão a 72 °C por 1 min e; d - uma etapa final de alongamento, a 72 °C por 10 minutos.

3) Exon 4 - tampão sulfato 1 X, 1,5 mM de cloreto de magnésio ($MgCl_2$), 0,1 mM de dNTP, 0,4 M de cada oligonucleotídeo iniciador, 0,5 U de Taq DNA polimerase (*Taq Platinum*, Invitrogen®) e 10 ng de DNA genômico, em um volume final de 10 μ l. As condições de ciclagem foram: a - desnaturação inicial a 96 °C por 5 minutos; b - 15 ciclos de desnaturação a 96 °C por 1min, hibridação a 66 °C por 30 seg e extensão a 72 °C por 45 seg; c - 30 ciclos de desnaturação a 96 °C por 1min, hibridação a 62 °C por 30 seg e extensão a 72 °C por 45 seg e; d - uma etapa final de alongamento, a 72 °C por 10 minutos.

4) Exon 5 - tampão sulfato 1 X, 2 mM de cloreto de magnésio ($MgCl_2$), 0,2 mM de dNTP, 0,2 M de cada oligonucleotídeo iniciador, 0,75 U de Taq DNA polimerase (*Taq Platinum*, Invitrogen®) e 30 ng de DNA genômico, em um volume final de 10 μ l. As condições de ciclagem foram: a - desnaturação inicial a 96 °C por 5 minutos; b - 15 ciclos de desnaturação a 96 °C por 1min, hibridação a 62 °C por 45 seg e extensão a 72 °C por 2 min; c - 30 ciclos de desnaturação a 96 °C por 1min, hibridação a 58 °C por 45 seg e extensão a 72 °C por 2 min e; d - uma etapa final de alongamento, a 72 °C por 10 minutos.

5) Exon 7 e 9 - tampão sulfato 1 X, 2,5 mM de cloreto de magnésio ($MgCl_2$), 0,1 mM de dNTP, 0,4 M de cada oligonucleotídeo iniciador, 0,5 U de Taq DNA polimerase

(*Taq Platinum*, Invitrogen®) e 20 ng de DNA genômico, em um volume final de 10 μ l. As condições de ciclagem foram: a - desnaturação inicial a 94 °C por 3 minutos; b - 40 ciclos de desnaturação a 94 °C por 30 seg, hibridação a 59 °C por 30 seg e extensão a 72 °C por 1min e 30 seg e; c - uma etapa final de alongamento, a 72 °C por 10 minutos.

Após a amplificação, a qualidade da reação foi verificada através da corrida eletroforética a 70 V por 60 minutos em gel de agarose 1% utilizando como corante GelRed (Biotium®) 2 X em DMSO diluído em solução azul de bromofenol e visualizado sob luz UV. Os tamanhos dos fragmentos amplificados para os exons 1-2, três, quatro, cinco, 7-9 foram 2582 pb, 1059 pb, 261 pb e 1912 pb e 829 pb, respectivamente.

Os produtos de PCR a serem sequenciados foram purificados com as enzimas Exonuclease I (EXO I - USB®) e Fosfatase Alcalina de Camarão (SAP - USB®) para eliminar os reagentes em excesso e tornar possível a reação de sequenciamento. Foram utilizadas 4 U de EXO I e 0,7 U de SAP, para 8 μ l de solução de PCR. A reação de purificação é feita através da incubação a 37 °C por 45 min, seguindo-se 15 min a 80 °C para inativação das enzimas. Algumas amostras foram purificadas através do protocolo baseado na precipitação do DNA amplificado por PCR com Polietilenoglicol (PEG), etanol, baixa temperatura e centrifugações. Inicialmente, o produto de cada reação de PCR é transferido para um tubo 0,5 ml e é adicionado o mesmo volume de PEG 8000 20% NaCl 2,5 M (50 μ l). Esta mistura permanece em banho-maria a 37 °C por 30 minutos e é então centrifugada a 13000 rpm (*revolutions per minute*) por 20 minutos. O sobrenadante é descartado invertendo o tubo bruscamente. Adiciona-se 125 μ l de etanol 80% gelado. Após um minuto a solução é novamente centrifugada a 13000 rpm durante dois minutos e o sobrenadante descartado, virando o tubo bruscamente. Em seguida é adicionado 125 μ l de etanol absoluto gelado e o sobrenadante descartado. Na última etapa o tubo é posto para secar em banho seco e em seguida são adicionados 10 μ l de água ultrapura. A solução resultante é incubada em banho-maria a 37 °C até o dia seguinte, para ressuspensão.

A quantificação dos produtos de PCR é feita através de eletroforese em gel de agarose 1% com 2 μ l do marcador de quantificação (*Low DNA Mass Ladder*, Invitrogen®), por comparação com as bandas do marcador, após corrida eletroforética com corante GelRed (Biotium®) 2 X a 80 V por 60 minutos.

As reações de sequenciamento utilizaram diferentes quantidades de produto de PCR e de MIX para o sequenciamento (*DYEnamic ET Terminator Cycle*

Sequencing Kit, GE Healthcare®), dependendo do tamanho do fragmento a ser sequenciado. Para fragmento de até 300 pb são utilizados cerca 50ng de DNA e 2 µl de MIX de sequenciamento. Já para os fragmentos maiores são utilizadas 150-200ng de DNA e 3 µl de MIX. A quantidade de oligonucleotídeo iniciador para a reação de sequenciamento foi de 5 pMoles por reação. O volume final da reação foi ajustado para 10 µl. A reação tem um passo inicial de desnaturação a 95 °C por 3 minutos, seguida por 35 ciclos de desnaturação por 20 segundos a 95 °C, hibridação a 50 °C por 15 segundos e extensão a 60 °C por 1 minuto.

Após a reação de sequenciamento, uma nova purificação é realizada antes da injeção das amostras no sequenciador. Para a precipitação do DNA é adicionado 1 µl de acetato de amônio 7,5 M e 30 µl de etanol absoluto, seguindo-se a centrifugação a 4000 rpm por 45 minutos a 15 °C. A placa é então virada bruscamente para descartar o sobrenadante e centrifugada invertida rapidamente a 300 rpm. São adicionados 100 µl de etanol 70 % à temperatura ambiente e é feita uma centrifugação rápida a 3000 rpm com a placa aberta. O sobrenadante é novamente descartado virando bruscamente e a placa é centrifugada invertida a 300 rpm rapidamente. As amostras são ressuspensas em 10 µl de solução desnaturante *loading solution* (formamida, do kit de sequenciamento da GE Healthcare®) e armazenadas a -20 °C.

As corridas de sequenciamento foram realizadas no sequenciador MegaBACE® 1000. Os parâmetros de injeção e corrida foram variáveis entre os fragmentos sequenciados, dependendo da qualidade da reação de PCR e do tamanho do fragmento a ser sequenciado. A visualização dos cromatogramas e a análise da qualidade das sequências foi feita com o auxílio dos programas Phred, Phrap e Consed (EWING e GREEN, 1998; GORDON, ABAJIAN e GREEN, 1998; GREEN, 1994), em ambiente Linux. O alinhamento das sequências foi realizado em ambiente Windows através do programa ClustalW (THOMPSON, HIGGINS e GIBSON, 1994), disponível no BioEdit (HALL, 1999), que também foi utilizado para a editoração das sequências.

9.3 ANÁLISE DOS DADOS

9.3.1 Diversidade alélica, haplotípica e genotípica

A identificação da fase gamética de dois ou mais sítios heterozigotos encontrados e determinação dos genótipos de cada indivíduo foi realizada com o auxílio do programa *online dbLRC (Leukocyte Receptor Complex database)* disponível no NCBI (<<http://www.ncbi.nlm.nih.gov/>>). Para a identificação dos alelos que compõem o genótipo de cada indivíduo, o programa utiliza um banco de dados com as sequências de todos os alelos já descritos para o gene *KIR3DL1/S1*, e faz permutações entre as sequências informadas até obter o genótipo mais provável para aquele indivíduo (menor número de *mismatches*). O programa leva em consideração que pelo menos um dos alelos do indivíduo é um alelo já conhecido.

As frequências alélicas e genotípicas foram determinadas por contagem direta e as frequências nas distintas populações foram comparadas pelo teste exato de diferenciação populacional baseado nas frequências dos alelos, com o auxílio do pacote de programas ARLEQUIN® versão 3.11 (EXCOFFIER et al., 2005).

O valor de desequilíbrio de ligação padronizado pelo seu valor máximo possível, dadas as frequências alélicas, (D') e o coeficiente de correlação (r^2) entre os sítios polimórficos em cada população foram determinados utilizando o programa HAPLOVIEW v.4.1 (BARRETT et al., 2005).

As distâncias genéticas entre as populações foram estimadas através da distância de Cavalli-Sforza e Edwards (CAVALLI-SFORZA e EDWARDS, 1967), a partir das frequências alélicas. A partir das distâncias genéticas foi construído um dendrograma que mostra a relação entre as populações deste estudo e, outro, incluindo outras populações cujos dados foram retirados da literatura. Para construção dos dendrogramas foi utilizado o pacote de programas PHYLIP (FELSENSTEIN, 2004), empregando o método *Neighbor-Joining* (SAITOU e NEI, 1987). Para visualização dos dendrogramas foi utilizado o programa TreeView 1.6.6 (PAGE, 1996). As frequências dos motivos de ligação a HLA nas populações deste estudo também foram utilizadas para construção de um dendrograma, utilizando os mesmos programas e métodos que para as frequências alélicas.

O Programa NETWORK v.4.0.6 (<http://www.fluxus-technology.com/sharenet.htm>) foi utilizado para a construção da rede mutacional mínima entre os alelos *KIR3DL1/S1* identificados no conjunto das populações deste estudo, empregando o algoritmo Median Joining (MJ) (BANDELT, FORSTER e ROHL, 1999). Este mesmo programa foi utilizado para a construção da rede mutacional entre os motivos de ligação às moléculas HLA.

9.3.2 Análises evolutivas

A contagem dos sítios polimórficos ao longo das sequências identificadas nas populações deste estudo foi realizada com o auxílio do pacote de programas ARLEQUIN, que foi utilizado também para o cálculo do índice de fixação F_{st} por sítio polimórfico e da média do valor de F_{st} sobre todos os sítios.

Com o auxílio do programa MEGA v. 4.0.2 (TAMURA et al., 2007) foi construído um dendrograma para os alelos já descritos e os novos alelos identificados neste estudo para evidenciar a relação entre eles. Os testes de neutralidade seletiva D de Tajima (D'), teste D de Fu e Li sem grupo externo (D^*), teste F de Fu e Li sem grupo externo (F^*), teste H de Fay e Wu com grupo externo (H) e a diversidade nucleotídica (π) foram estimados com o auxílio do programa DnaSP v.5.10 (ROZAS et al., 2003), utilizando a análise de coalescência com 10000 repetições, para determinação da significância dos valores observados. A sequência homóloga *KIR3DL1/2* nos Chimpanzés (genebank: **NC_006486**) foi utilizada como grupo externo para o cálculo do teste H de Fay e Wu. O teste de homozigosidade de Ewens Watterson (θ_W) foi realizado através do programa ARLEQUIN.

A contagem de substituições não sinônimas por sítio não sinônimo (dN) e de substituições sinônimas por sítio sinônimo (dS) também foi realizada com o auxílio do programa DnaSP. Para os exons um, dois e sete, a proporção dN/dS foi verificada através do teste exato de Fisher, devido ao tamanho reduzido das sequências. Já para a sequência completa do gene *KIR3DL1/S1* e também para os exons 3, 4 e 5 do conjunto total de alelos e separadamente para os alelos *3DL1* e *3DS1*, a comparação foi feita através do teste *Codon Based Z-test* empregando o método de Nei-Gojobori, considerando o modelo de substituição de Jukes & Cantor (25% de probabilidade para cada nucleotídeo), com o auxílio do programa MEGA (v. 4.0.2). Esta mesma

abordagem foi empregada para verificar a proporção dN/dS nos códons específicos dos exons quatro e cinco que codificam os motivos de ligação à HLA.

A possibilidade de coevolução entre *KIR3DL1/S1* e *HLA* foi verificada através da observação de coocorrência dos genótipos *KIR3DL1/S1* e dos motivos *Bw4* e *Bw6* nas moléculas HLA, comparando a ocorrência conjunta observada e esperada (calculada a partir dos produtos das frequências observadas em cada população). Para estas comparações foram utilizadas tabelas de contingência MxN, com o auxílio do programa RxC (MILLER, 1997). A correlação entre as frequências das linhagens *3DL1* e *3DS1* e as moléculas HLA com epítipo *Bw4* e *Bw6* foi verificada através do teste de correlação de postos de Spearman (<http://www.spearmansrank.com/>) e da correlação de Pearson (programa Bio Estat v. 5.0).